March 1950

BIOMETRICS

Vol. 6 No. 1

THE BIOMETRICS SECTION, AMERICAN STATISTICAL ASSOCIATION

A Procedure for Quantifying Subjective Appraisals of Odor, Flavor and Texture of Foodstuffs — *J. W. Hopkins*

The Significance of Deviations from Expectation in a Poisson Series — *R. A. Fisher*

The Estimation of Error in Rectangular Lattices — *P. M. Grundy*

The General Theory of Prime-Power Lattice Designs, V — *W. T. Federer*

An Application of Sequential Analysis to Whitefish Sampling — *G. B. Oakland*

Proceedings of the Second International Biometric Conference

- The Present Status of Biometry — *W. G. Cochran*
- Teaching and Education in Biometry — *M. S. Bartlett*

The Biometrics Section of the
American Statistical Association

BIOMETRICS

TABLE OF CONTENTS

A Procedure for Quantifying Subjective Appraisals of Odor, Flavor and Texture of Foodstuffs . . . . J. W. HOPKINS 1

The Significance of Deviations from Expectation in a Poisson Series . . . . . R. A. FISHER 17

The Estimation of Error in Rectangular Lattices, P. M. GRUNDY 25

The General Theory of Prime-Power Lattice Design, V W. T. FEDERER 34

An Application of Sequential Analysis to Whitefish Sampling G. B. OAKLAND 59

Proceedings of the Second International Biometric Conference 69

The Present Status of Biometry . . . . . W. G. COCHRAN 75

Teaching and Education in Biometry . . M. S. BARTLETT 85

Queries 99

Abstracts 103

Number 1 March 1950 Volume 6

Material for *Biometrics* should be addressed to the Chairman of the Editorial Board, Institute of Statistics, North Carolina State College, Raleigh, N. C.; and material for Queries should go to "Queries", Statistical Laboratory, Iowa State College, Ames, Iowa, or to any member of the committee.

As the consideration of articles for publication in *Biometrics* is dependent upon referees, it is necessary for us to have *three* copies of each paper.

We are taking this opportunity to request that in the future all manuscripts be submitted in triplicate.

GERTRUDE M. COX, *Editor*

Officers
American Statistical Association

President, Samuel S. Wilks; *President-elect*, Lowell J. Reed; *Past President*, Simon Kuznets; *Directors*, Gertrude M. Cox, W. Edwards Deming, Cyril H. Goulden, Frederick F. Stephan, Willard L. Thorp, Louis L. Thurstone; *Vice-Presidents*, Dorothy S. Brady, Harold A. Freeman, Philip M. Hauser; *Secretary-Treasurer*, Samuel Weiss; *Council*, Charles M. Armstrong, Waite S. Brush, W. G. Cochran, Donald R. G. Cowan, M. I. Gershenson, Morris H. Hansen, Howard L. Jones, Thomas J. Mills, Paul R. Rider, David Schneider, John R. Stockton, Eliot J. Swan, John Tukey, W. Allen Wallis, Sylvia C. Weyl.

Officers
Biometrics Section

Chairman, Harold F. Dorn; *Vice-Chairman*, Leonard J. Savage; *Secretary*, Jerome Cornfield; *Section Committee*, Frank Wilcoxon, George W. Snedecor, Frederick Mosteller, Dorothy J. Morrow, Margaret Merrell, Lila Knudsen, John W. Fertig, A. E. Brandt, Joseph Berkson, Joseph Zubin.

Editorial Board
Biometrics

Chairman, Gertrude M. Cox; *Members*, C. I. Bliss, W. G. Cochran, Churchill Eisenhart, J. W. Fertig, H. C. Fryer, Horace Norton, A. M. Mood, G. W. Snedecor, and Jane Worcester.

Membership dues in the American Statistical Association are as follows: regular members $8.00, contributing members $25.00, student members (for a period of 4 years) $4.00, family membership (an additional $2.00 for husband and wife receiving only one copy of the periodicals) $2.00, introductory dues for persons under 30 (for the first year only) $4.00, associate members of the Biometrics Section $4.00, senior members (for persons over 60 having been members of the ASA for 25 years or more) exempt, life membership as voted on October 7, 1948, i.e., 25 to 30 years $200.00, over 30 at Age A $290.00 – 3A.

Annual subscriptions are as follows: To the Journal of *The American Statistical Association*, for non-members $8.00; allocation of members dues $5.00; rate to members of associated or affiliated societies $5.00; to *Biometrics* $4.00, *Biometrics* for American Statistical Association Members $3.00; to *The American Statistician* $1.50. Subscription and applications should be sent to the American Statistical Association, 1603 K Street, N. W., Washington 6, D.C.

Entered as second-class matter, May 25, 1945, at the post office at Washington, D.C., under the Act of March 3, 1879. Biometrics is published four times a year—in March, June, September and December—by the American Statistical Association for its Biometrics Section. Editorial Office: 1603 K Street, N.W., Washington 6, D.C.

A PROCEDURE FOR QUANTIFYING SUBJECTIVE APPRAISALS OF ODOR, FLAVOR AND TEXTURE OF FOODSTUFFS*

J. W. HOPKINS†

National Research Council, Ottawa

ABSTRACT

A SCORING PROCEDURE assigning the integers 0 to 5 to six defined gradations ranging from "none" to "gross" of intensity of abnormal or undesirable attributes and those from −5 to +5 to eleven gradations ranging from "gross deficiency" to "gross excess" of typical or desirable attributes having a preferred intensity is described, and from experimental data reported it is concluded that this may reasonably be expected to provide results of approximately calculable reproducibility related in a consistent way to the composition of test materials. In tests with unselected groups of some 20 individuals, mean scores recorded increased in proportion to relative increments in single sensory stimuli and up to 8 test samples could be appraised for two or three attributes concurrently. Both inter-individual differences in discrimination and intra-individual inconsistencies contributed to the observed variability in single scores, and on the 0 to 5 scale especially the relation of variance to mean tended in practice to resemble that of a binomial variate.

INTRODUCTION

Experimentation in food technology often requires small-scale subjective "consumer" appraisals of the flavor, odor or texture of a series of test materials, and a numerical expression of these appraisals for correlation with other data.

Subjective appraisals may be quantified by ordinally ranking an entire test series or by scoring each sample individually. Ranking, a well-established psychometric technique used successfully for example by Cameron (2) in his studies of relative sweetness, has advantages when the test samples are few in number or can be grouped without restriction in small sets in accordance with the requirements of "incomplete block" or "lattice" experimental arrangements (8, 13). This how-

*Contribution from the Division of Applied Biology, National Research Laboratories, Ottawa. Paper No. 4 on Applied Mathematical Statistics, No. 239 of the Canadian Committee on Food Preservation, and N.R.C. No. 2075.

†Biometrician

ever may be impossible in extensive processing or storage experiments. Moreover even a coarse scoring system, if effective, may be more informative than a ranking and hence be preferred by investigators.

One method of scoring the intensity of abnormal, unpleasant or undesirable attributes may be arrived at by defining six gradations of subjective appraisal corresponding to the integers 0 to 5 as follows:

Appraisal	*Score*
none	0
slight	1
moderate	2
decided	3
very decided	4
gross	5

Similarly if the intensity of characteristic or desirable attributes having an optimal or preferred level is in question, eleven gradations of appraisal with corresponding scores of -5 to $+5$ may be defined thus:

Appraisal	*Score*
gross deficiency	-5
very decided deficiency	-4
decided deficiency	-3
moderate deficiency	-2
slight deficiency	-1
ideal level	0
slight excess	$+1$
moderate excess	$+2$
decided excess	$+3$
very decided excess	$+4$
gross excess	$+5$

Some statistical characteristics of scores obtained by this method in the subjective appraisals unit of the Food Investigation Laboratories, Division of Applied Biology, National Research Council of Canada are reported upon below.

EXPERIMENTAL RESULTS

Relation of Scores to Test Stimuli

For a test of discrimination, eight concentrations of each of three odoriferous substances were incorporated in fresh commercial shortening (hydrogenated cooking fat). The resulting 24 test materials were then each scored for intensity of undesirable or unpleasant odors twice in random groupings of 2, 4 and 8, i.e. six times in all, by 21 individuals,

11 female and 10 male, recruited from the scientific, technical and administrative staff of the laboratories but unselected in respect of ability or previous experience, and untrained except for instruction in a few items of simple standardized procedure. Scoring was done by each individual independently in the privacy of a partitioned appraisal laboratory, with the results shown graphically in Fig. 1, in which each

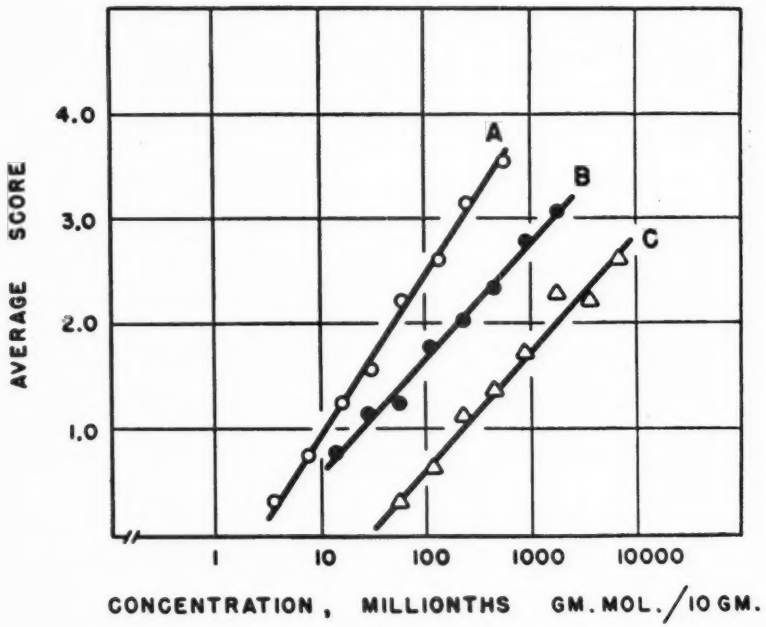

FIGURE 1.

MEANS OF 21 SEXTUPLICATE SCORES IN RELATION TO CONCENTRATION OF ODORIFEROUS TEST SUBSTANCES. A, METHYL HEXYL KETONE (PUNGENT); B, *n*-BUTYL ALDEHYDE (SWEETISH-PUNGENT); C, *n*-CAPRYLIC ACID (RANCID). ABSCISSAL SCALE IS LOGARITHMIC.

plotted point represents a mean of 126 scores and the horizontal scale is logarithmic. It is evident that over the range dealt with the average scores for all three test substances, namely methyl hexyl ketone (pungent), *n*-butyl aldehyde (sweetish-pungent) and *n*-caprylic acid (rancid) tended to increase in proportion to relative increments in concentration, this type of reaction being, as noted by Dermer (7), also characteristic of taste as well as of other sensations. The average score for methyl

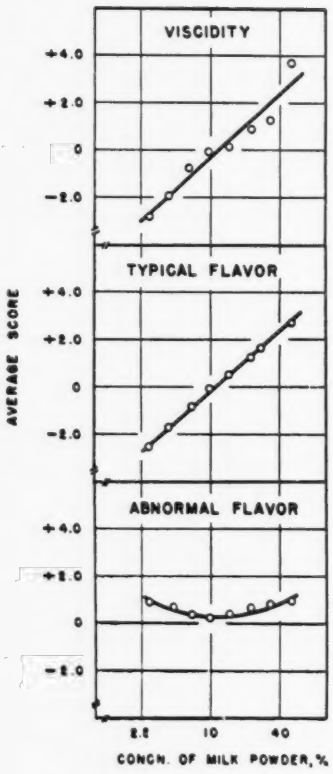

FIGURE 2A.

MEANS OF 23 SEXTUPLICATE SCORES FOR VISCIDITY AND INTENSITY OF TYPICAL AND ABNORMAL MILK FLAVOR IN RELATION TO CONCENTRATION OF MILK POWDER IN AQUEOUS SUSPENSIONS. ABSCISSAL SCALE IS LOGARITHMIC.

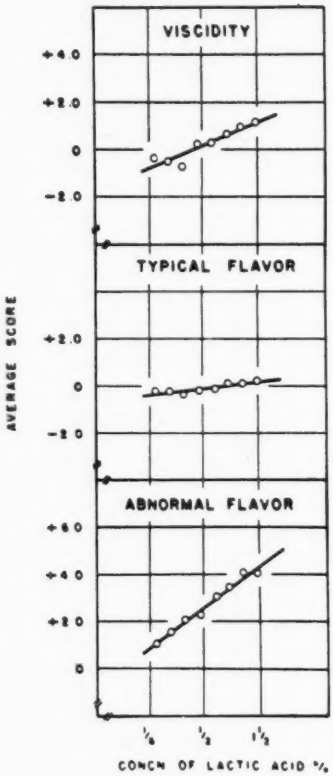

FIGURE 2B.

MEANS OF 23 SEXTUPLICATE SCORES FOR VISCIDITY AND INTENSITY OF TYPICAL AND ABNORMAL MILK FLAVOR IN RELATION TO CONCENTRATION OF LACTIC ACID IN AQUEOUS SUSPENSIONS CONTAINING 10% MILK POWDER. ABSCISSAL SCALE IS LOGARITHMIC.

hexyl ketone increased by 1.6 and that for the other two substances by 1.1 units for each 10-fold increase in concentration.

Milk powder, actually a commercial preparation of skim milk fortified with lactose, and lactic acid provided sixteen test materials, eight containing different amounts of milk powder suspended in water and eight having a fixed amount of milk powder with graded additions of lactic acid, for a similar test of taste discrimination. Each of these materials

TABLE I

REGRESSION EQUATIONS RELATING AVERAGE SCORE (S) TO CONCENTRATION OF KNOWN STIMULI (X)

Attribute subjectively scored	Known variable stimulus		Regression equation
	Description	Experimental range and unit	
Intensity of abnormal odor	Conc. of methyl hexyl ketone	4 to 500 millionths gm. mol./10 gm.	$S = 1.55 \log X - 0.46$
Intensity of abnormal odor	Conc. of *n*-butyl aldehyde	14 to 1776 millionths gm. mol./10 gm.	$S = 1.10 \log X - 0.49$
Intensity of abnormal odor	Conc. of *n*-caprylic acid	55 to 7103 millionths gm. mol./10 gm.	$S = 1.12 \log X - 1.60$
Viscidity of fluid	Conc. of milk powder in water	2.9 to 50.0 per cent	$S = 4.53 \log X - 4.87$
Viscidity of fluid	Conc. of lactic acid	0.19 to 1.43 per cent	$S = 2.07 \log X - 1.37$
Intensity of typical "milk" flavor	Conc. of milk powder in water	2.9 to 50.0 per cent	$S = 3.99 \log X - 4.22$
Intensity of typical "milk" flavor	Conc. of lactic acid	0.19 to 1.43 per cent	$S = 0.62 \log X - 0.55$
Intensity of abnormal flavor in milk	Conc. of milk powder in water	2.9 to 50.0 per cent	$S = 1.61 (\log X)^2 - 3.35 \log X + 2.08$
Intensity of abnormal flavor in milk	Conc. of lactic acid	0.19 to 1.43 per cent	$S = 3.68 \log X - 0.13$

was tasted, again twice in random groupings of 2, 4 and 8, by 23 individuals, 11 female and 12 male, of whom 5 females and 4 males had taken part in the foregoing odor test, and scored for viscidity ("body"), intensity of typical "milk" flavor and intensity of abnormal or undesirable flavor. The results for samples with and without lactic acid are shown separately in Fig. 2a and 2b, each plotted point here representing the mean of 138 scores. Irregularities are manifest, but may be reasonably attributed to factors such as instability of some of the more concentrated or acidulated suspensions, and in general the scores recorded were undoubtedly related in a consistent way to the composition of the test samples. Some overlapping of presumed sensations of typical and abnormal flavor is to be noted, but a significant differentiation between these attributes was nevertheless achieved. Table I summarizes in the form of weighted (see below) regression equations the average relations between score and composition resulting from both the foregoing odor and flavor tests.

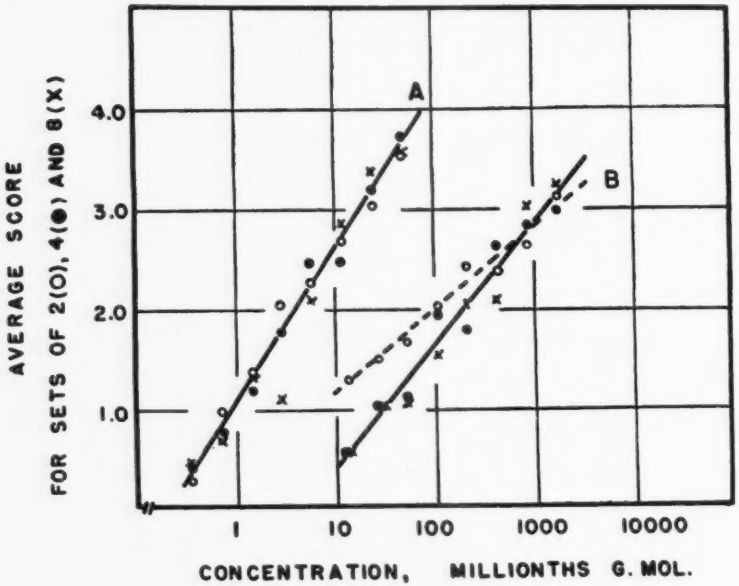

FIGURE 3.

MEANS OF 21 DUPLICATE SCORES IN RELATION TO CONCENTRATION OF ODORIFEROUS TEST SUBSTANCES AND TO NUMBER OF SAMPLES APPRAISED CONCURRENTLY. A, METHYL HEXYL KETONE; B, n-BUTYL ALDEHYDE. ABSCISSAL SCALE IS LOGARITHMIC.

No consistent difference in the scores recorded by males and females was demonstrable in either test, and varying the number of samples submitted for concurrent appraisal significantly affected the relation of average score to composition in only one instance. As illustrated at "*B*" in Fig. 3, the subjects tested were more sensitive to differences in the intensity of *n*-butyl aldehyde odor when samples were submitted for appraisal in sets of 4 or 8 than when they were submitted in pairs.

Impressions of odor or flavor attributes of most foodstuffs will of course result from compound stimuli of unknown intensity (10). However Fig. 4, in which each plotted point represents the mean of only

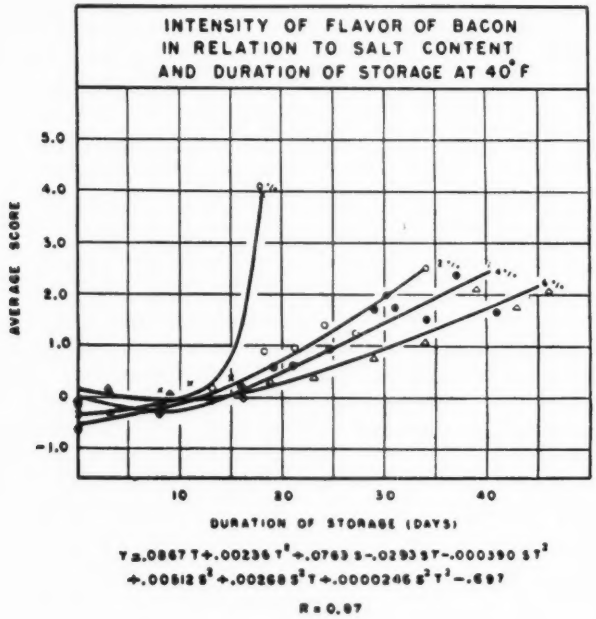

FIGURE 4.

MEANS OF 23 SINGLE SCORES FOR INTENSITY OF FLAVOR OF CANADIAN WILTSHIRE BACON IN RELATION TO SALT CONTENT AND DURATION OF STORAGE AT 40°F. DATA FROM EXPERIMENT 4(c) OF TABLE II.

23 individual scores, indicates that such impressions when quantified by the foregoing procedure may at least exhibit a reasonable degree of mutual consistency.

Reproducibility of Scores

Variability in recorded scores may result from differences in sensory threshold or in discriminatory sensitivity of individuals (1, 10) as well as from random inconsistencies. If therefore p individuals each score q test materials on r occasions, the single score S_{ijk} given by the ith individual to the jth sample at the kth repetition may conveniently be regarded as a sum of additive components

$$S_{ijk} = A + S'_{i..} + S'_{.j.} + S'_{ij.} + S'_{ijk}$$

of which A denotes the general mean for the trial as a whole; $S'_{i..}$ the deviation from A of the mean $S_{i..}$ of all qr scores recorded by the ith individual; $S'_{.j.}$ the deviation from A of $S_{.j.}$, the mean of all pr scores given to the jth test material; $S'_{ij.}$ the amount by which the deviation from $S_{.j.}$ of $S_{ij.}$, the mean of the r scores given to the jth material by the ith individual, differs from $S'_{i..}$; and S'_{ijk} the deviation from $S_{ij.}$ of S_{ijk}.

The variance of S'_{ijk}, of $S'_{ij.}$ and of $S'_{i..}$, or of combinations of them, actually occurring in any particular experiment is determinable from the p rows, q columns and pq cells each containing r items, of an analysis of variance table of the pqr individual scores S_{ijk}. Inconsistencies in the scoring of the same material by the same individual on different occasions will then be directly reflected in the observed variance of S'_{ijk}. The observed variance of $S'_{ij.}$ will, however, also be generated in part by these intra-individual inconsistencies as well as by any underlying consistent discrepancies in the relative scores assigned to specific materials in the same test series by different individuals, and the net component of variance ascribable to the latter must therefore be estimated in terms of an analysis of variance table (6), from [Mean sq. $(S'_{ij.})$ − Mean sq. $(S'_{ijk})]/r$. In the same way the recorded variance of $S'_{i..}$ will be generated in part by both the foregoing components of variation, together with any inequalities in the average score given to the series of test materials as a whole by different individuals, and the net variance actually corresponding to this last must accordingly again be estimated, in terms of an analysis of variance table, from [Mean sq. $(S'_{i..})$ − Mean sq. $(S'_{ij.})]/rq$. Table II lists the net components of variance, deduced in this way, ascribable to $S'_{i..}$, to $(S'_{ij.} + S'_{ijk})$ and to S'_{ijk} in each of nine experiments participated in by different groups of individuals recruited from the previously mentioned classes of laboratory staff. In all of these experiments some 20 degrees of freedom were available for estimation of the variance of $S'_{i..}$ and from 109 to 1176 d.f. for estimation of each of the other values given. Experiment 7 extended over a period of two weeks, Experiment 3 over four and Experiments 4 and 5 over twenty weeks each. The conclusions to be

TABLE II

COMPONENTS OF AVERAGE INTER- AND INTRA-INDIVIDUAL VARIANCE OF RECORDED SINGLE SCORES

Experiment No.	Attribute scored for intensity	Scoring scale	Mean score	Average net variance of		
				$S'_{i..}$	$S'_{ij.} + S'_{ijk}$	S'_{ijk}
1 (a)	Odor of methyl hexyl ketone	0 to +5	+1.8	0.18	0.57	0.56
2 (a)	Viscidity of milk powder suspensions	−5 to +5	0.0	0.11	0.85	0.61
2 (b)	Sourness (lactic acid) in same	0 to +5	+0.6	0.43	0.94	0.83
2 (c)	Typical "milk" flavor in same	−5 to +5	+0.1	0.09	1.40	0.78
3 (a)	Abnormal odor in processed whole milk	0 to +5	+0.7	0.13	0.43	—
3 (b)	Abnormal flavor in same	0 to +5	+1.2	0.28	0.55	—
4 (a)	Texture of processed and stored bacon	−5 to +5	+0.2	0.37	0.82	—
4 (b)	Abnormal odor in same	0 to +5	+1.3	0.29	0.86	—
4 (c)	Typical "bacon" flavor in same	−5 to +5	+0.7	0.27	1.64	—
5	Abnormal flavor in processed and stored bacon	0 to +5	+0.7	0.06	0.84	—
6	Abnormal odor in processed and stored poultry	0 to +5	+0.7	0.20	0.52	—
7 (a)	Texture of processed and stored butter	−5 to +5	+0.4	0.43	1.60	—
7 (b)	Abnormal odor in same	0 to +5	+1.1	0.12	0.68	—
7 (c)	Abnormal flavor in same	0 to +5	+1.1	0.34	0.85	—
8 (a)	Viscidity of processed edible oils	−5 to +5	−0.2	0.54	0.35	—
8 (b)	Typical odor in same	−5 to +5	+0.1	0.53	0.63	—
8 (c)	Typical flavor in same	−5 to +5	+0.6	0.51	1.10	—
9 (a)	Abnormal flavor in preserved fruit	0 to +5	+0.7	0.22	0.76	—
9 (b)	Typical flavor in same	−5 to +5	−0.3	0.17	1.98	—

drawn from these results are that intra-individual inconsistencies were on the average the largest single source of variance in the scores recorded in the two experiments, Nos. 1 and 2, in which this component was investigated explicitly, but that in general the effect of one or both specified components of inter-individual variation was also appreciable. On the whole, appraisals of "typical" attributes made on the less restricted scale of −5 to +5 and involving preference as well as sensory reaction were more variable than those of "abnormal" attributes on the scale of 0 to 5, and appraisals of flavor more variable than those of odor and texture or viscidity.

Relation of Variance to Mean

The phrase "on the average" was used advisedly above, for when as here the range of single scores S_{ijk} is limited, the possible deviations $S'_{i..}$, $S'_{ij.}$ and S'_{ijk} are also necessarily restricted, and in particular $S'_{ij.}$ and S'_{ijk} for specified j (test material) cannot be mutually independent of the mean $S_{.j.}$.

When S_{ijk} is restricted to the integers 0 to 5 its total variance, i.e. var $(S'_{i..} + S'_{ij.} + S'_{ijk})$ must be zero when $S_{.j.}$ is either maximal (5.00) or minimal (0.00); and has been found in practice in these laboratories to be greatest when $S_{.j.}$ is intermediate. As $S_{.j.}$ ranges from 0.00 to 5.00 the associated variance of S_{ijk} thus tended to vary as $(S_{.j.})(5 - S_{.j.})$, and hence to be approximately equal to $\lambda(S_{.j.})(5 - S_{.j.})$, which expression then specifies an approximate scedastic curve of var $(S'_{i..} + S'_{ij.} + S'_{ijk})$ as a function of $S_{.j.}$. "Weighted least squares" estimates of coefficients of proportionality λ have therefore been computed from $\sum w_i$ var $(S'_{i..} + S'_{ij.} + S'_{ijk})(S_{.j.})(5 - S_{.j.})/\sum_i w_i [(S_{.j.})(5 - S_{.j.})]^2$. In this expression var $(S'_{i..} + S'_{ij.} + S'_{ijk})$ and $S_{.j.}$ are observed quantities, while the weights w_i should be inversely proportional to the sampling variance of the former. This is unknown but may be expected to tend to $[\sigma^2(S'_{i..} + S'_{ij.} + S'_{ijk})]^2/p$. Weights w'_i corresponding to the observed var $(S'_{i..} + S'_{ij.} + S'_{ijk})$ for each j were therefore used to compute first approximations λ'. The squares of the expected var $(S'_{i..} + S'_{ij.} + S'_{ijk})$ obtained from $\lambda'(S_{.j.})(5 - S_{.j.})$ were then used to calculate a second series of weights w_i , which were applied to the observations to arrive at a second approximation to λ. In this way, numerical values ranging from $\lambda = 0.13$ for Experiment 1 of Table II to $\lambda = 0.31$ for Experiment 5 have been obtained. The scedastic curve resulting from this latter is shown in relation to the observed variances in Fig. 5a.

An evident analogy with binomial variation (12, chap. 16) therefore suggested that although the transformation of scale $\phi_{ijk} = \sin^{-1}\sqrt{(1/5)(S_{ijk})}$ could not be expected completely to stabilize the total variance of single scores, the similar transformarion $\phi_{.j.} = \sin^{-1}\sqrt{(1/5)(S_{.j.})}$ might be quite effective in this regard when applied to the group means $S_{.j.}$ constituting the effective units of observation in experiments in which there is actual replication of treatments or in which second or higher order interaction mean squares are used in tests of significance. The results portrayed in Fig. 5b, obtained by transforming the mean scores $S_{.j.1}$ and $S_{.j.2}$ for two independent random subgroups of 12 of the 25 individuals appraising each of the 52 test materials represented in Fig. 5a, conform to this supposition.

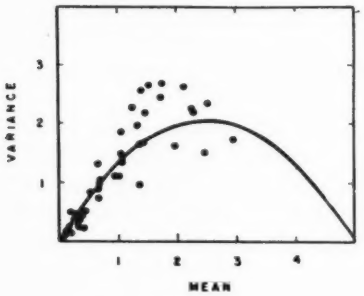

FIGURE 5A.

VARIANCE OF 25 INDIVIDUALS' SCORES FOR INTENSITY OF ABNORMAL FLAVOR ON 6-POINT SCALE IN RELATION TO MEAN. DATA FROM EXPERIMENT 5 OF TABLE II.

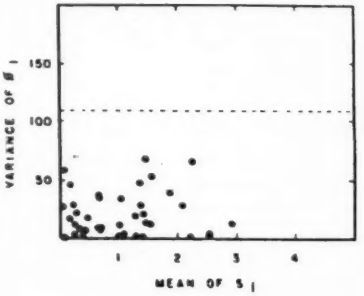

FIGURE 5B.

VARIANCE OF TRANSFORMED SUBMEANS $\phi_{.j.} = \text{SIN}^{-1}\sqrt{(1/5)(S_{.j.})}$ IN RELATION TO UNTRANSFORMED MEAN $S_{.j.}$ (DATA AS IN FIG. 5A). DASHED LINE INDICATES UPPER 2% POINT FOR THE VARIANCE OF A NORMAL DEVIATE HAVING THE SAME S.D.

In experiments in which the S_{ijk} may assume integral values from -5 through 0 to $+5$, the total variance of $(S'_{i..} + S'_{.j.} + S'_{ijk})$ must also be zero when $S_{.j.}$ is maximal or minimal. It has been found however that in at least some such instances this variance may for practical purposes be regarded as independent of $S_{.j.}$ in the central part of the latter's range, with which technologists are mainly concerned. This situation is exemplified in the results for Experiment 4(c) of Table II, shown in Fig. 6a. On the other hand an appreciable measure of agree-

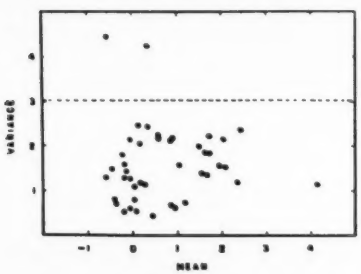

FIGURE 6A.

VARIANCE OF 23 INDIVIDUALS' SINGLE SCORES FOR INTENSITY OF BACON FLAVOR ON 11-POINT SCALE IN RELATION TO MEAN. DATA AS IN FIG. 4. DASHED LINE INDICATES UPPER 2% POINT FOR THE VARIANCE OF A NORMAL DEVIATE HAVING THE SAME S.D.

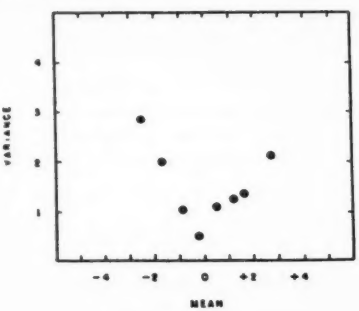

FIGURE 6B.

VARIANCE OF 23 INDIVIDUALS' SINGLE SCORES FOR INTENSITY OF MILK FLAVOR ON 11-POINT SCALE, ESTIMATED FROM DATA OF EXPERIMENT 2(C) OF TABLE II.

ment between individuals as to the "preferred" intensity of the attribute in question may cause the relation of variance to mean to tend towards

$$\text{Var.}\ (S'_{i..} + S'_{ij.} + S'_{ijk}) = \lambda(\mid S_{.j.} \mid + a)(5 - \mid S_{.j.} \mid)$$

illustrated in Fig. 6b by the results for samples without lactic acid in Experiment 2(c). As in this instance a was not unduly large, the transformation $\phi_{.j.} = \sin^{-1}\sqrt{(1/5)(S_{.j.})}$ again had a considerable stabilizing effect on the total variance of group mean scores. This situation however may be infrequent.

Relation of Variance to Some Other Factors

Table III lists some ratios of the total variance of $(S'_{i..} + S'_{ij.} + S'_{ijk})$ for samples appraised 4 and 8 at a time to that for samples appraised in pairs. Values given are the averages of those computed separately from the untransformed scores for each of the test materials in Experiments 1 and 2 of Table II. Here, reproducibility of individual ap-

TABLE III

AVERAGE RATIOS OF ESTIMATED VARIANCE (INTER- AND INTRA-INDIVIDUAL) OF SCORES FOR SAMPLES APPRAISED IN SETS OF 2, 4 AND 8

Experiment No.	Attribute scored for intensity	Scoring scale	Average ratio of est. var. $(S'_{i..} + S'_{ij.} + S'_{ijk})$	
			4:2	8:2
1(a)	Odor of methyl hexyl ketone	0 to +5	1.0	1.0
1(b)	Odor of *n*-butyl aldehyde	0 to +5	1.0	1.1
1(c)	Odor of *n*-caprylic acid	0 to +5	1.3	0.9
2(a)	Viscidity of milk powder—water suspensions	−5 to +5	1.9	1.7
2(b)	Sourness (lactic acid) in milk	0 to +5	0.8	0.9
2(c)	Characteristic flavor of milk	−5 to +5	1.2	1.4

praisals of viscidity was most affected by the number of samples scored concurrently. Even in this instance the increase in variance was less than proportional to the associated reduction in time required for appraisal of the test series in larger subsets, but these appraisals were for the most part simple in character, and Crocker (5, chap. 11) is of the opinion that no more than 8 food or beverage samples can be effectively rated concurrently in commercial testing.

TABLE IV

EFFECT OF INCREASING SOURNESS (LACTIC ACID) ON VARIANCE OF INDIVIDUAL APPRAISALS OF INTENSITY OF TYPICAL "MILK" FLAVOR

Lactic acid in test sample, %	Score for intensity of typical "milk" flavor	
	Mean	Variance
0.00	−0.1	0.47
0.19	−0.2	0.94
0.25	−0.2	1.73
0.33	−0.3	1.52
0.45	−0.2	5.56
0.60	−0.1	3.05
0.80	+0.1	6.74
1.07	+0.1	7.36
1.43	+0.2	6.50

Table IV shows the effect of increasing sourness on the mean $S'_{.j.}$ and total variance of $(S'_{i..} + S'_{ij.} + S'_{ijk})$ of individual scores for intensity of typical "milk" flavor of the eight test materials in Experiment 2(b) containing 10% milk powder together with varying amounts of lactic acid, and the one in Experiment 2(c) having the same amount of milk powder without any lactic acid. The mean scores $S'_{.j.}$ showed only a slight progressive bias specified in Table I and illustrated in the centre panel of Fig. 2b, but the disturbing effect of the second flavor component on appraisal of the first was reflected in a roughly 10-fold increase in the average total variance of single scores. In contrast, corresponding variances differed little when these nine test materials were appraised in random groupings of 2, 4, and 8 of the whole series of 16 test materials, those for groupings of 4 and 8 being in the average ratio of 0.8:1 and 1.0:1 respectively to those for pairs.

There is however some evidence suggesting that when a series of test samples is appraised in u sets of v each, the rating of a specified sample may sometimes be conditioned by the nature of the $v - 1$ other samples in the same set, an undesirable sample for example being considered more distasteful if encountered in a set in which all others are of good quality. In 16 such tests of the flavor of various foodstuffs it was arranged that while all of the p individuals participating tasted all uv samples, each of k independent random groupings of the test samples into u sets of v was submitted to $n = p/k$ individuals. It was therefore possible to distinguish, in analyses of variance of the transformed ap-

praisal scores, between the discrepance $S'_{ij.}$ of "individuals × samples" of individuals appraising the same, and that of individuals appraising different random groupings. In 5 of these 16 tests the latter exceeded the former by more than the conventional 0.05 probability level, whereas in the absence of any real effect, the expected frequency of such an excess would (assuming normality) be 0.8. When the residual discrepancies ("error variances") are heterogeneous, the larger must be used to test the significance of differences between means. A necessary condition for the statistical validity of such tests of significance is that the groupings of the samples into sets be independently made in an objectively random manner, so that the component of discrepance resulting from grouping is a chance variable.

No consistent difference in the variance of scores recorded by male and female subjects was noted in these tests.

DISCUSSION

These results suggest that a scoring system of the type considered might be used for appraisal of odor, flavor or physical attributes of varying degrees of generality or specificity with a reasonable expectation of results of approximately calculable reproducibility, related in a consistent way to the composition of test materials. The findings would of course be strictly valid only for the local sub-population (in our case Ottawa middle-class urban) of which the experimental subjects were representative. Differences in preference levels have been reported for example not only between regions (11) but between different age-groups in the same district of the U.S.A. (9), and it is with such subjective reactions, whether or not associated with actual sensory differences, that the food technologist is concerned.

In this restricted sense, observed internal variances may be used to compute approximate "fiducial limits" or "necessary differences" for group mean scores. Although the sampling distribution of recorded group means $S_{.j.}$ must become progressively skewed in the vicinity of either end of the range, that of random differences between pairs of such means will on a "null hypothesis" of an identical parent population remain symmetrical. "Normal" tests of significance of differences using the appropriate variance, although only approximate are therefore unlikely to be grossly misleading even in inter-group comparisons, if the groups involved are of about the same size and include as here 20 or more individuals.

Inter-group reproducibility of absolute mean scores will be affected by all the three components of inter- and intra-individual variation

denoted above by $S'_{i..}$, $S'_{ij.}$, and S'_{ijk} , whereas only the last two of these will affect inter-group reproducibility of intra-group *differences*, the $S'_{i..}$ operating as random variables in the first instance and as fixed effects in the second. Accordingly it may be calculated for example from the values given in Table II for Experiment 4(c) (intensity of typical "bacon" flavor), in which mean and variance were independent over the range in question, that the "normal" approximation to the difference between two mean scores each recorded by a 24-member group necessary for the customary 5% level of statistical significance would be in the neighborhood of 0.8 gradations of appraisal for inter-group, and of 0.7 for intra-group comparisons. When the variance of $S'_{ij.}$ is not zero, the inter-group variance of intra-group comparisons will be minimal for a fixed number n of appraisals per sample when n individuals each make a single appraisal. This procedure, adopted in Experiments 3–9 of Table II, will also minimize the variance of inter-group comparisons if the variance of either $S'_{i..}$ or $S'_{ij.}$ is significant. In view of the possibility of discrepancies due to grouping when a test series must be appraised in subsets, it is also desirable to maximize the number of chance combinations by making an independent random grouping for each appraiser.

When the variance is related to the mean in a manner resembling that of a binomial variate, the inverse sine transformation may sometimes render the data of extensive experiments more amenable to analysis of variance or covariance techniques by facilitating tests of significance while still providing efficient estimates of main effects (4). It is however difficult to define an interaction effect on this scale in a physically meaningful way (3). Preference may therefore be given to a weighted regression analysis of group mean scores from factorial experiments, such as Expts. 4(b) and 5 above, in which the factors occur at 4, 6 or more levels, some treatment effects are large, and the object is to estimate the magnitude, under specified conditions, of effects whose reality has previously been demonstrated. Appropriate weights, inversely proportional to the expectation of $(S_{.j.})(5 - S_{.j.})$, may be estimated from a graphical first approximation, but the necessity of estimating and applying them constitutes the major statistical inconvenience of the appraisal procedure.

ACKNOWLEDGMENT

The foregoing subjective appraisals were all made under the technical supervision of Elinor M. Hamilton, who also assisted in the statistical examination of the resulting scores.

REFERENCES

1. Blakeslee, A. F. and Salmon, T. N. Genetics of sensory thresholds: individual taste-reactions for different substances. *Proc. Nat. Acad. Sci.* (Washington), 21, 84–90, 1935.
2. Cameron, A. T. The taste sense and the relative sweetness of sugars and other sweet substances. Sugar Research Foundation Inc. New York, Scientific Report No. 9, 1947.
3. Cochran, W. G. Some difficulties in the statistical analysis of replicated experiments. *Empire J. Exp. Agric.* 6, 157–175, 1938.
4. Cochran, W. G. Analysis of variance when experimental errors follow the Poisson or binomial laws. *Annals Math. Statistics*, 11, 335–347, 1940.
5. Crocker, E. C. *Flavor*. McGraw-Hill, New York, 1945.
6. Crump, S. L. The estimation of variance components in analysis of variance. *Biometrics Bull.* 2, 7–11, 1946.
7. Dermer, O. C. The science of taste. *Proc. Okla. Acad. Sci.* 27, 9–20, 1947.
8. Fisher, R. A. *The design of experiments*. 4th ed. Oliver and Boyd, London, 1947.
9. Harding, P. O. Teen-age students versus adults as taste judges of Temple oranges. *Food Research* 13, 6–10, 1948.
10. Hopkins, J. W. Precision of assessment of palatability of foodstuffs by laboratory panels. *Can. J. Research*, F, 24, 203–214, 1946.
11. Quartermaster Food and Container Institute for the Armed Forces Committee on Food Research. Food Acceptance Research. Quartermaster Corps Manual QMC 17–9. War Department, Office of the Quartermaster General, Washington, 1946.
12. Statistical Research Group, Columbia University *Techniques of statistical analysis*. McGraw Hill, New York, 1947.
13. Yates, F. The design and analysis of factorial experiments. Imperial Bureau of Soil Science, Harpenden, England, Technical Communication No. 35, 1937.

THE SIGNIFICANCE OF DEVIATIONS FROM EXPECTATION IN A POISSON SERIES

R. A. Fisher

University of Cambridge

In a series of observations of whole numbers in which the number r is observed a_r times, we may set

$$S_r(a_r) = A,$$

$$S_r(ra_r) = B,$$

then the ratio

$$\bar{r} = B/A$$

is a sufficient estimate of m, where m is the parameter of the Poisson Series with expectations

$$m_r = A\, e^{-m} \frac{m^r}{r!}.$$

If $\bar{r}$ is small and the series short, evidence of deviation may come chiefly or wholly from frequencies with low expectation, so that the measure of general discrepancy

$$\chi^2 = S_r(a_r - m_r)^2/m_r$$

will have a sampling distribution not accurately given by the usual table of χ^2. Equally, the special test for discrepancy of the variance, the index of dispersion

$$\chi^2 = S_r\{a_r(r - \bar{r})^2\}/\bar{r}$$

for $A - 1$ degrees of freedom, will be suspect owing to the low estimated expectation, $\bar{r}$, in each of the A cells.

PARTICULAR EXAMPLE

These remarks may be illustrated by the particular example

r	frequency
0	124
1	12
2	2
3	2
Total	140

The mean, $\bar{r}$, is only 11/70; the index of dispersion comes to $219.\dot{8}\dot{1}$. If we ignore the recommendation that χ^2 should only be relied on if the expectation in each cell exceeds 5, and apply it to a case in which the expectation is only 0.157, we have

	$\sqrt{2\chi^2}$	20.9675
$\sqrt{2n-1}$ =	$\sqrt{277}$	16.6433
Difference		4.3242

The difference indicates strong significance, at a level of about one in a hundred thousand ($P < .00001$), but the experimenter is still left in some doubt as to its validity.

The conventional application of the more comprehensive χ^2 test is at least equally embarrassing. Using the sufficient estimate, $m = \bar{r}$, we have

	observed	expected	ungrouped	grouped
r	a	m_r	$(a - m_r)^2/m_r$	$(a - m_r)^2/m_r$
0	124	119.6415	.1588	.1588
1	12	18.8008	2.4600	2.4600
2	2	1.4772	.1850	
3	2	.0774	47.76	3.8293
>3	0	.0031	.0031	
Total	140	140.0000	50.57	6.4481
			$n = 3$	$n = 1$

The χ^2 test with three degrees of freedom suggests high significance, owing to the two observations of the variate 3, but the expectation in this class is obviously too low for the test to be valid. The test with one degree of freedom shows significance just beyond the 1% point.

Some, however, will object that the expectation in the third class is still only 1.5577, while others may feel that the test has been made unduly insensitive by pooling the twos with the threes, thereby ignoring the strong evidence which the latter really provide.

THE EXACT FREQUENCY DISTRIBUTION

In a rigorous treatment of the problem we must distinguish between two different facets of it. (i) The calculation of an exact frequency distribution giving the objective frequencies of the observational series we have obtained, and of other alternative observational series. (ii) The choice of a criterion by which the assemblage of possibilities is to be divided up in making a test of significance. The first facet is purely mathematical; only one answer is possible. The second takes account of the needs and prior ideas of the experimenter; by it he specifies the particular question he chooses to ask. The two procedures illustrated in Section 2 attempted answers to two different questions: (a) Is the sample more variable than a sample with this mean should be, and if so at what level of significance; (b) Do the frequencies observed differ significantly from those expected in a Poisson Series, having the same mean, and if so at what level of significance. With an exact distribution other questions may suggest themselves.

The key to the distributional problem is that the only parameter of the Poisson Series is capable of sufficient estimation. From this it follows that the Likelihood function of m is identical for all observational series having the same size and the same mean $\bar{r}$. The relative frequencies of all such series are therefore entirely independent of the unknown parameter. These frequencies can be calculated. There will, it is true, be a great many of them equal indeed to the number of partitions of the partible number 22, but we shall not need to calculate all of these, especially of the rarer, in order to answer such questions as: How frequently is

$$S\{a_r(r - \bar{r})^2\}/\bar{r}$$

less than $219.\dot{8}\dot{1}$, or, more expeditiously, how often is

$$S_r(a_r r^2)$$

less than 38. Or, if we consider the second test, to answer the question: How frequently is

$$S_r(a_r - m_r)^2/m_r$$

less than 50.57. If either of these are the questions of our choice, they can clearly be answered.

If, for any value m of the unknown parameter

$$p_r = e^{-m} m^r / r!,$$

then the probability of an observational sequence a_0 , a_1 , a_2 , $\cdots$ is

$$\frac{A!}{a_0!\, a_1!\, a_2! \cdots} p_0^{a_0} p_1^{a_1} p_2^{a_2} \cdots$$

Whence, substituting for p_r , we have the constant function of m

$$e^{-Am} m^B$$

with the particular factor of the distribution, which is independent of m,

$$\frac{A!}{a_0!\, a_1!\, a_2! \cdots} \cdot \frac{1}{(2!)^{a_2} (3!)^{a_3} \cdots}.$$

Moreover the sum of all such probabilities will be the probability of observing the number B in a Poisson series with a parameter Am, and is

$$e^{-Am} \frac{(Am)^B}{B!}.$$

Consequently, the relative frequency of any observable series, having A and B as specified, will be

$$\frac{B!}{A^B} \cdot \frac{A!}{a_0!\, a_1!\, a_2! \cdots} \cdot \frac{1}{(2!)^{a_2} (3!)^{a_3} \cdots},$$

and this we can calculate without knowledge of m.

A LIST OF CONFIGURATIONS

The possible observed distributions having $A = 140$, $B = 22$ correspond to the partitions of the partible number 22. If A were less than B, we should also stipulate that the number of parts should not exceed A. The correspondence may be shown in a short table

Partition	a_0	a_1	a_2	a_3
(1^{22})	118	22	—	—
(21^{20})	119	20	1	—
$(2^2 1^{18})$	120	18	2	—
(31^{19})	120	19	—	1

and so on. Corresponding with any distribution we can now list the corresponding frequencies, and any criterion value the use of which may be in view, as in Table 1.

The list of the first forty-five partitions in order of frequency extends, further than is needed for any ordinary criteria. The reason for this is that using a variety of tests, different criteria are liable to put the distributions in different orders, so that when treated exactly, slightly different questions receive slightly different answers.

The first point to be noted is that the probability of observing exactly the series we have observed is .000222; consequently on no view is the level of significance higher than one in 4500. It should not, however, be placed so high as this, since on any reasonable view some other possible series will be judged more discrepant than the one we have observed.

Next, the sum of the probabilities of the series observed together with all less frequent series is .000639. This gives a possible test of significance at a level of about one in 1600. It is not a good test for general use, but seems to be not inappropriate for the present problem. The kind of case in which it fails to carry conviction is the not uncommon one in which it is a matter of arbitrary choice whether a class of observational series, which may be distinguished, but which seem to differ in nothing relevant to the actual problem under discussion, shall be considered separately or together; for this choice will decide its order relative to other classes of observation. Mere order of frequency cannot therefore in general be taken to provide an order of discrepancy.

The standard large-sample χ^2 tests obviously provide valid, but, for our purpose, somewhat conventional tests. The index of dispersion depends on the integer $S(r^2a_r)$, which for our sample is 38. The probability of observing a value 38 or higher is .001112, at a level of about one in 900. But the probability of observing actually 38 is .0007355, and the fact that this is a large fraction of the total, while not invalidating the test of significance, shows that it is somewhat less sensitive than it might be, for a reason extraneous to our purpose, namely, that it is somewhat heavily grouped. Of course, if our purpose is sharply defined as "to test if the sample variance is too high for a Poisson series, and if so at what level of significance," this limitation is inherent in the problem we have set ourselves.

For the general χ^2 test the probability of χ^2 as great as or greater than 50.499 is .002684, or rather more than one in 400. In comparison with the previous result this shows that though on either test the significance of the deviation is high, yet the deviations are particularly

TABLE 1

Partition	Frequency	$S(r^2a_r)$	χ^2	$S(a \log (ar!)) - 273.222$
(21^{20})	.340167	24	0.315	0.090
(2^21^{18})	.269299	26	0.301	0.079
(1^{22})	.175238	22	2.125	0.793
(2^31^{16})	.113506	28	2.083	0.295
(2^41^{14})	.027911	30	5.661	0.972
(321^{17})	.026707	30	11.346	0.792
(31^{19})	.018898	28	12.485	1.354
(32^21^{15})	.014886	32	12.004	0.938
(32^31^{13})	.004236	34	14.458	1.430
(2^51^{12})	.004130	32	11.036	1.787
(421^{16})	.000930	36	324.66	2.261
(41^{18})	.000742	34	328.57	2.770
(32^41^{11})	.000666	36	18.708	2.208
(3^31^{16})	.000620	34	49.716	2.738
(3^221^{14})	.000605	36	49.249	2.342
(42^21^{14})	.000454	38	325.84	2.466
(2^61^{10})	.000366	34	18.207	2.837
$(3^22^21^{12})$	.000222	38	50.499	2.675
(42^31^{13})	.000111	40	331.23	3.027
(32^51^9)	.000059 per million	38	24.754	3.258
(431^{15})	40.34	40	340.31	3.486
$(3^22^21^{10})$	39.08	40	53.705	3.385
(4321^{13})	34.16	42	340.07	3.303
(521^{15})	24.29	44	10466.	3.958
(51^{17})	21.97	42	10467.	4.311

of a kind to increase the variance, rather than in other ways to increase the general value of χ^2. The contrast is heightened by the fact that the general χ^2 is not grouped.

It will be seen, however, that the conventional χ^2 value is absurdly sensitive to high observations such as 4 and 5, and that such occurrences govern its value. A better general test when expectations are liable to be very small is provided by

$$\underset{r}{S}\, a_r \log \frac{a_r}{m_r},$$

which, if the logarithms are to the base e, agrees with

$$\tfrac{1}{2}\chi^2$$

TABLE 1—*Continued*

Partition	Frequency	$S(r^2a_r)$	χ^2	$S(a \log (ar!)) - 273.222$
$(2^7 1^8)$	18.84	36	27.173	4.139
$(42^1 1^{10})$	14.66	42		
$(432^1 1^{11})$	10.66	44		
$(52^1 1^{13})$	10.29	46		
$(3^1 1^{13})$	7.59	40		
$(3^2 21^{11})$	4.74	42		
$(3^2 2^1 1^8)$	3.49	42		
$(32^2 1^7)$	2.79	40	32.597	
$(52^1 1^{11})$	2.14	48		
$(432^1 1^9)$	1.55	46		
$(42^2 1^8)$	1.05	44		
$(3^2 2^2 1^3)$	1.03	44		
(531^{14})	.98	48		
(5321^{12})	.71	50		
$(4^3 1^{14})$	.61	46		
$(43^2 1^{13})$	.59	46		
$(2^8 1^6)$	.50	38	37.936	
(61^{16})	.50	52		
(621^{14})	.49	54		
$(4^2 21^{12})$	.44	48		

when the values of m are large enough for the latter to be valid.

For convenience I have taken logarithms to the base 10, and use in the last column the easily calculable criterion

$$S_r (a_r \log_{10} (a_r r!)) - 273.222;$$

my values would therefore have to be multiplied by 2 log. 10 or 4.6 to give values comparable to χ^2.

This measure of discrepancy also is not grouped. The value for our sample is 2.675 and the probability of a value as great as this or greater is .002367, at a rather higher level of significance than the large sample χ^2 test when treated exactly.

SUMMARY

These tests make it clear that the sample departs from expectation based on the Poisson Series at a level, for a general test, of about $\frac{1}{4}\%$; while a corresponding refinement of the Index of Dispersion, the measure

of variance relative to the mean, claims a level of about one in 900, in spite of the fact that this measure, owing to the grouping of different samples, is inherently somewhat insensitive. It is in the direction of increased dispersion that our sample departs most distinctly from the Poisson expectation.

The use is illustrated of a generalised measure of deviation,

$$S\left\{a \log\left(\frac{a}{m}\right)\right\},$$

which is essentially the logarithmic difference in likelihood between the most likely Poisson Series and the most likely theoretical series without restriction. This measure seems well fitted to take the place of the conventional χ^2, when class expectations are small.

FIGURE 1

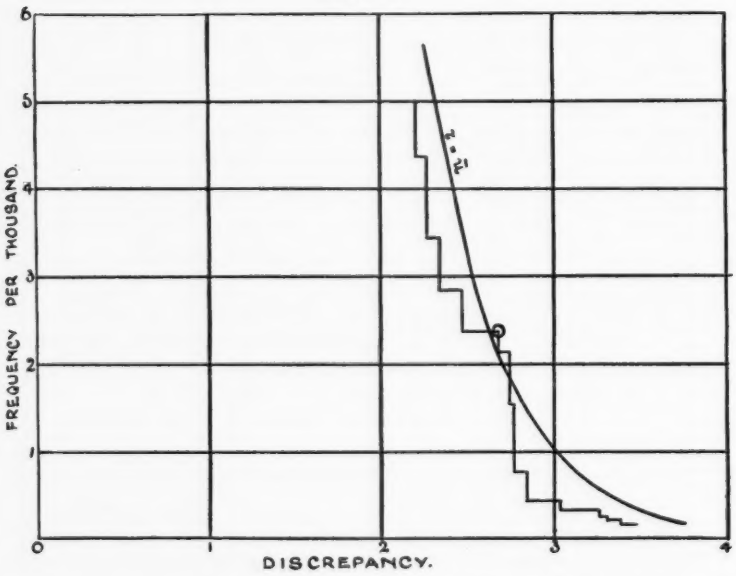

Figure 1 compares the actual discontinuous distribution of this criterion with that of the theoretical χ^2 for two degrees of freedom.

THE ESTIMATION OF ERROR IN RECTANGULAR LATTICES

P. M. GRUNDY

Rothamsted Experimental Station

INTRODUCTION

THE RECTANGULAR LATTICES considered in this note are the designs for agricultural variety trials, and similar experiments, discussed by Harshbarger briefly in [1] and more fully in [2]. The analysis of variance involves determination of two components (a) and (b) of the sum of squares for blocks eliminating varieties (and replications), and of the within-block error sum of squares. In [2] the within- and between-block error variances are estimated by somewhat elaborate linear combinations of these three sums of squares. It would of course be simpler (after computing the analysis of variance as before) to use, besides the within-block error sum of squares, merely the sum of Components (a) and (b). The object of this note is to show that the simpler method is generally better.

The design is for $p(p+1)$ varieties ($p \geq 2$) in incomplete blocks of p plots. The varieties are allotted to the $p(p+1)$ ordered pairs of unequal integers (i, j) of Fig. 1:

	$(1, 2)$	$(1, 3)$	$\cdots$ $(1, p)$	$(1, p+1)$
$(2, 1)$		$(2, 3)$	$\cdots$ $(2, p)$	$(2, p+1)$
. . . . .				
$(p+1, 1)$	$(p+1, 2)$	$(p+1, 3)$	$\cdots$ $(p+1, p)$	

FIGURE 1.

There are (say) r "x-replications", in each of which the rows of Fig. 1 are taken as blocks, and r "y-replications", in which the columns of Fig. 1 give the blocks. In the field, plots are randomized within blocks, and blocks within replications. In [2] the value of r was restricted to 2, but a mere pair of replications ($r = 1$) may be used when the number of varieties is very large.

In the statistical analysis, the plot-yields per acre are taken to be the sums of the respective varietal constants, replication constants, plot contributions, and block contributions. The different plot and block contributions are assumed to be independent normal deviates, with variances σ^2 for the former, τ^2 for the latter.

Notation

The use of an arbitrary r necessitates some departures from the notation of [2]. It should also be noted that the preceding description of the experiment differs formally from [2], and follows instead the more convenient arrangement adopted in [3].

Consider first the varieties in the i'th row of Fig. 1; we write B_{xik} for their total in the k'th x-replication, B_{xi} for their total over all x-replications, and T_{yi} for their total over all y-replications. Similarly B_{yik} , B_{yj} , and T_{xj} refer to the varieties in the j'th column of Fig. 1. Thus the B_{xik} and B_{yik} are block totals, and

$$B_{xi} = \sum_{k=1}^{r} B_{xik} , \qquad B_{yi} = \sum_{k=1}^{r} B_{yik} .$$

The total of the k'th x-replication is denoted by R_{xk} , and the total of all x-replications by S_x ; similarly R_{yk} and S_y refer to the y-replications. V_{ij} , where $i \neq j$, denotes the total of the variety (i, j); and m_{ij} denotes the estimate of mean yield of that variety, based on intra- and inter-block information. In the rest of this paper, the suffixes i, j, k are always to be given all values allowed by the definitions. The summation sign will apply to such of i, j, k as occur, but not to x and y, which are used descriptively.

Analysis

Expressions for the varietal estimates m_{ij} when $r = 2$ were given in [2], and amended formulae for the variances of their differences in [3]. We take the opportunity here to record the corresponding results for general r. The formulae involve the weights w, w' appropriate to within- and between-block comparisons, which are given by

$$w = 1/\sigma^2, \qquad w' = 1/(\sigma^2 + p\tau^2).$$

It is preferable, however, to eliminate w' (or τ^2) by the introduction of a ratio γ, defined by

$$\gamma = \frac{w - w'}{w + w'} = \frac{p\tau^2}{2\sigma^2 + p\tau^2} \qquad \text{.....(1)}$$

The expression for m_{ij} is then

$$m_{ij} = \frac{1}{2r}(V_{ij} + C_{xi} + C_{yj}) \quad \cdots\cdots\cdots\cdots\cdots\cdots (2)$$

where the C_{xi} and C_{yj} (whose values should be tabulated in the appropriate places) are given by

$$\left.\begin{aligned} C_{xi} &= \frac{\gamma}{p^2 - \gamma^2}\{-p(B_{xi} - T_{yi}) + \gamma(B_{yi} - T_{xi})\} \\ C_{yj} &= \frac{\gamma}{p^2 - \gamma^2}\{-p(B_{yj} - T_{xj}) + \gamma(B_{xj} - T_{yj})\} \end{aligned}\right\} \quad \cdots\cdots (3)$$

In relation to the present design, there are four types of varietal differences $m_{ij} - m_{kl}$, each with a different variance. These types (of which the second does not occur in the trivial case $p = 2$) are: (a) varieties in the same block, so that $i = k$ or $j = l$; (b) i, j, k, l all different; (c) $i = l$ or $j = k$ but not both; (d) $i = l$ and $j = k$. The typical formulae for the variances are:

$$\text{var}\,(m_{12} - m_{13}) = \frac{\sigma^2}{r}\left\{1 + \frac{p\gamma}{p^2 - \gamma^2}\right\} \quad \cdots\cdots\cdots\cdots (4)$$

$$\text{var}\,(m_{12} - m_{34}) = \frac{\sigma^2}{r}\left\{1 + \frac{2p\gamma}{p^2 - \gamma^2}\right\} \quad \cdots\cdots\cdots\cdots (5)$$

$$\text{var}\,(m_{12} - m_{23}) = \frac{\sigma^2}{r}\left\{1 + \frac{2p\gamma + \gamma^2}{p^2 - \gamma^2}\right\} \quad \cdots\cdots\cdots\cdots (6)$$

$$\text{var}\,(m_{12} - m_{21}) = \frac{\sigma^2}{r}\left\{1 + \frac{2p\gamma + 2\gamma^2}{p^2 - \gamma^2}\right\} \quad \cdots\cdots\cdots\cdots (7)$$

The analysis of variance for general r, obtainable by methods similar to those in [2], takes the form shown in Fig. 2 and equations (8)–(10).

$$\text{Component (a)} = \frac{1}{p}\left(\sum B_{xik}^2 + \sum B_{yik}^2\right) - \frac{1}{rp}\left(\sum B_{xi}^2 + \sum B_{yi}^2\right)$$

$$- \frac{1}{p(p+1)}\left(\sum R_{xk}^2 + \sum R_{yk}^2\right) + \frac{1}{rp(p+1)}(S_x^2 + S_y^2) \quad \cdots\cdots (8)$$

$$\text{Component (b)} = \frac{1}{2r(p^2-1)}\{p\sum (B_{xi} - T_{yi})^2 + p\sum (B_{yi} - T_{xi})^2$$

$$- 2\sum (B_{xi} - T_{yi})(B_{yi} - T_{xi}) - 2(S_x - S_y)^2\} \quad \cdots\cdots\cdots (9)$$

	Sums of Squares	D.F.
Replications	$\frac{\sum R_{xk}^2 + \sum R_{yk}^2}{p(p+1)} - \frac{(S_x + S_y)^2}{2rp(p+1)}$	$2r - 1$
Varieties alone	$\frac{\sum V_{ij}^2}{2r} - \frac{(S_x + S_y)^2}{2rp(p+1)}$	$p^2 + p - 1$
Blocks, eliminating varieties	Component (a), from (8) or (10)	$2(r-1)p$
	Component (b), from (9)	$2p$
		$2rp$
Intra-block error	By subtraction	$(2r-1)(p^2-1) - p$
Total	Total sum of squares	$2rp(p+1) - 1$

FIGURE 2.

When $r = 1$, Component (a) vanishes; when $r = 2$, equation (8) reduces to

$$\text{Component (a)} = \frac{1}{2p}\{\sum (B_{xi1} - B_{xi2})^2 + \sum (B_{yi1} - B_{yi2})^2\}$$

$$- \frac{1}{2p(p+1)}\{(R_{x1} - R_{x2})^2 + (R_{y1} - R_{y2})^2\} \cdots\cdots\cdots\cdots (10)$$

Moments of the Mean Squares

In conformity with [2], the mean squares derived from Component (a), Component (b), and from the intra-block error sum of squares will be denoted by N, P, Q respectively. Various linear combinations of these could be used as estimates of σ^2 and τ^2. We shall use the criterion of variance for the comparison of such estimates, and for that purpose require not only the expectations but also the second-order moments of N, P, Q.

In the first place, the joint distribution of N and P is independent of Q, since N and P are expressible as functions of the block and variety totals, and these are uncorrelated with the deviations which contribute to Q. Moreover, the differences $B_{xi} - T_{yi}$ and $B_{yi} - T_{xi}$ which occur in (9) are easily seen to be uncorrelated with the interaction contrasts

of the B_{xik} , and likewise of the B_{yik} . Since (8) presents Component (a) as the sum of squares for these two sets of interactions, it follows that N and P are independent. *Thus N, P, Q are independent variates.*

The expectation of Q is of course

$$\mathrm{E}(Q) = \sigma^2 \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots (11)$$

The B_{xik} are independent normal variates with variance $p(\sigma^2 + p\tau^2)$, and the mean of each is the sum of parameters depending respectively on i (block-type) and k (replications). The interaction sum of squares for this two-way classification is therefore distributed as $p(\sigma^2 + p\tau^2) \cdot \chi^2$. Similarly for the y-replications. Thus N is proportional to a χ^2 variate; and, allowing for the factor $1/p$ in (8), we also have

$$\mathrm{E}(N) = \sigma^2 + p\tau^2 \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots (12)$$

Since N and Q are proportional to χ^2 variates, their variances are given by the squares of their means, divided by half the respective numbers of degrees of freedom. Hence

$$\left.\begin{aligned} \mathrm{var}\,(N) &= \frac{(\sigma^2 + p\tau^2)^2}{(r-1)p} \\ \mathrm{var}\,(Q) &= \frac{2\sigma^4}{(2r-1)(p^2-1) - p} \end{aligned}\right\} \cdots\cdots\cdots\cdots (13)$$

It is a feature of the analysis of variance that the contrasts involved in Component (b) are orthogonal to replication totals. Consequently the moments of P may, without loss of generality, be calculated on the assumption that all the replication constants are equal. In that case the normal variates $(B_{xi} - T_{yi})$, etc., in the equation

$$\left.\begin{aligned} 4rp(p^2-1)P = p \sum (B_{xi} - T_{yi})^2 + p \sum (B_{yi} - T_{xi})^2 \\ - 2 \sum (B_{xi} - T_{yi})(B_{yi} - T_{xi}) - 2(S_x - S_y)^2 \end{aligned}\right\} \cdots\cdots (14)$$

are reduced to normal *deviates*. The relevant variances and covariances are given in equations (16) and (17), and thence it is found (cf. [2]) that

$$\mathrm{E}(P) = \sigma^2 + \tfrac{1}{2}p\tau^2 \cdots\cdots\cdots\cdots\cdots\cdots\cdots (15)$$

The variance of P is more troublesome, and for this calculation it is convenient to write

$$K = \sigma^2 + p\tau^2, \qquad L = \tau^2.$$

In terms of these quantities, the variances of the $(B_{xi} - T_{yi})$ etc. are given by

$$\left.\begin{aligned} \text{var}\,(B_{xi} - T_{yi}) &= rp(2K - (p-1)L) \\ \text{var}\,(S_x - S_y) &= 2rp(p+1)K \end{aligned}\right\} \cdots\cdots\cdots (16)$$

and their covariances by

$$\text{cov}\,(B_{xi} - T_{yi}\,,\, B_{yi} - T_{xi}) = 0 \cdots\cdots\cdots\cdots\cdots\cdots\cdots (17)$$

$$\left.\begin{aligned} \text{cov}\,(B_{xi} - T_{yi}\,,\, B_{yj} - T_{xj}) &= -2rK \qquad (i \neq j) \\ \text{cov}\,(B_{xi} - T_{yi}\,,\, B_{xj} - T_{yj}) &= r(p-1)L \qquad (i \neq j) \\ \text{cov}\,(B_{xi} - T_{yi}\,,\, S_x - S_y) &= 2rpK \end{aligned}\right\} \cdots (18)$$

together with the similar equations obtained by interchanging x and y.

Using equations (16)–(18), we now evaluate var (P) with the help of the lemma that *if* $u_1\,, \cdots, u_4$ *are jointly normal deviates,*

$$\text{cov}\,(u_1u_2\,,\, u_3u_4) = \text{cov}\,(u_1\,,\, u_3)\,\text{cov}\,(u_2\,,\, u_4) + \text{cov}\,(u_1\,,\, u_4)\,\text{cov}\,(u_2\,,\, u_3).$$

(Consideration of the moment generating function gives a short proof of this.) It follows that, in the absence of replication effects,

$$\begin{aligned} &\text{var}\,(\textstyle\sum (B_{xi} - T_{yi})^2) \\ &\quad = (p+1)\cdot 2r^2p^2(2K - (p-1)L)^2 + p(p+1)\cdot 2r^2(p-1)^2L^2 \\ &\quad = 8r^2p^2(p+1)K^2 - 8r^2p^2(p^2-1)KL + 2r^2p(p^2-1)^2L^2. \end{aligned}$$

Similarly, in the absence of replication effects,

$$\begin{aligned} &\text{var}\,(\textstyle\sum (B_{xi} - T_{yi})(B_{yi} - T_{xi})) \\ &\quad = 4r^2p(p+1)^2K^2 - 4r^2p^2(p^2-1)KL + r^2p(p^2-1)^2L^2 \end{aligned}$$

and

$$\text{var}\,((S_x - S_y)^2) = 8r^2p^2(p+1)^2K^2.$$

Further applications of the lemma give, in the absence of replication effects, the covariances of the quadratic sums in equation (14):

$$\text{cov}\,(\textstyle\sum (B_{xi} - T_{yi})^2,\ \sum (B_{yi} - T_{xi})^2) = 8r^2p(p+1)K^2,$$

$$\operatorname{cov}\left(\sum (B_{xi} - T_{yi})^2, \sum (B_{xi} - T_{yi})(B_{yi} - T_{xi})\right)$$

$$= -4r^2p(p^2 - 1)KL,$$

$$\operatorname{cov}\left(\sum (B_{xi} - T_{yi})^2, (S_x - S_y)^2\right) = 8r^2p^2(p + 1)K^2,$$

$$\operatorname{cov}\left(\sum (B_{xi} - T_{yi})(B_{yi} - T_{xi}), (S_x - S_y)^2\right) = -8r^2p^2(p + 1)K^2.$$

The variance of P is now derived from (14) by a straightforward calculation. We obtain

$$\operatorname{var}(P) = \frac{1}{p} K^2 - KL + \frac{(p^2 + 1)}{4p} L^2,$$

that is

$$\operatorname{var}(P) = \frac{1}{p}\left(\sigma^2 + \frac{1}{2} p\tau^2\right)^2 + \frac{1}{4p} \tau^4 \quad \cdots\cdots\cdots\cdots (19)$$

As a matter of fact, it could have been predicted that var (P) would be expressible in the form*

$$\frac{1}{p}\left(\sigma^2 + \frac{1}{2} p\tau^2\right)^2 + c\tau^4$$

with some c depending only on p and r. The formula could then have been determined completely by working out var (P) with a convenient value given to σ^2/τ^2.

Estimation of σ^2 and τ^2

In investigating estimates of σ^2 and τ^2, we can confine attention to unbiased linear combinations of N, P and Q. It is now possible to compare the variances of such estimates.

The independent unbiased estimates of σ^2 under consideration are Q and $2P - N$, and by (13) and (19) the ratio of their variances is as

$$\frac{2p(r - 1)(1 - \gamma)^2}{(1 + \gamma)^2 + 4(r - 1)(p^2 + \gamma^2)/p^2} \quad \text{to} \quad (2r - 1)(p^2 - 1) - p,$$

where γ is defined by (1). Even when $\gamma = 0$, this ratio is never greater than 2/35 over the whole range $p \geq 3$, $r \geq 1$. Thus there is no appreciable gain in taking a weighted mean of Q and $2P - N$, instead

*To see that, consider the $2p$ linear combinations of the yields involved in Component (b). A system of coordinates $z_1, \cdots, z_{2p}$ for these can be chosen so that the z_i are independent variates with variances $\sigma^2 + \lambda_i\tau^2$ respectively, where the λ_i depend only on p and r. Then $2pP = z^2_1 + \cdots + z^2_{2p}$, and the statement will be verified if $\mathrm{E}(P)$ and $\operatorname{var}(P)$ are worked out in terms of p, σ^2, τ^2, and the λ_i.

of Q. In practice we have $\gamma > 0$, and then $2P - N$ becomes relatively even more erratic, although the need for a tolerably accurate estimate of σ^2 is in no way diminished. In the case $r = 2$, $p = 5$, the weighted mean suggested by Harshbarger [2] actually has a larger variance than Q when $w/w' \geq 6$; yet the numerical illustration there gives $w/w' = 10$, apart from errors of estimation.

Having settled on Q as estimate of σ^2, we have still to choose a suitable estimate $\hat{\tau}^2$ of τ^2. The appropriate variance criterion* now appears to be not the crude variance of $\hat{\tau}^2$, but its residual variance after elimination of linear regression on Q. This residual variance is minimized by combining N and $2P$ with weights inversely proportional to their variances; that is, with weights bearing the ratio of

$$4(r-1)(p^2+\gamma^2) \qquad \text{to} \qquad p^2(1+\gamma)^2.$$

If a value of w/w' near 1 were expected, this would lead to the employment of

$$2(r-1)N + P,$$

which may be compared with [2], p. 12. On the other hand, a value of w/w' near 3 would make $4(p^2+\gamma^2)/p^2(1+\gamma)^2$ near 2, which would lead us to work with

$$(r-1)N + P.$$

Further increase in w/w' increases the relative importance of P, but decreases the need for accurate estimation of τ^2. As a compromise, the last-mentioned combination of N and P appears to be most appropriate; and it has the added advantage of being, apart from a factor $2p$, simply the sum of squares for blocks eliminating varieties. The present design is obviously intended for experimental material in which $w/w' > 1$, and in agricultural field trials 3 is quite a moderate value for that ratio.

The expectation of the mean square for blocks eliminating varieties is, from (12) and (15),

$$\frac{1}{r}\mathrm{E}((r-1)N + P) = \sigma^2 + \left(1 - \frac{1}{2r}\right)p\tau^2.$$

---

*The possibility of estimating σ^2 with variance less than that of Q is, as has been shown, of no practical importance. However, this circumstance presents us with a dilemma: minimising the variance of the estimate of τ^2 is not the same as minimising that of $\sigma^2 + p\tau^2$. This is avoided by using the residual variance, which moreover seems most appropriate when viewed geometrically.

Conclusion

It has been shown that the following formulae are eminently suitable for the estimation of σ^2 and τ^2:

$$\hat{\sigma}^2 = Q,$$

$$\hat{\sigma}^2 + \left(1 - \frac{1}{2r}\right)p\hat{\tau}^2 = \text{mean square for blocks eliminating varieties}$$

$$= \frac{1}{2rp}\,(\text{Component (a)} + \text{Component (b)}).$$

These results were obtained by consideration of minimum variance.

This work was carried out during the tenure of a post-graduate scholarship granted by the Ministry of Agriculture.

REFERENCES

[1] Harshbarger, B. Preliminary report on the rectangular lattices. *Biometrics* 2, 115–119, 1946.

[2] Harshbarger, B. Rectangular lattices. *Virginia Ag. Exp. St. Memoirs* 1, 1–26, 1947.

[3] Harshbarger, B. Correction to [2], to appear in *Biometrics*.

Correction, Volume 5, Number 4 of *Biometrics* in "The Choice of a Response Metameter in Bio-Assay" by D. J. Finney.

Page 268 two lines from bottom substitute $i < -1$ for $i < 1$

THE GENERAL THEORY OF PRIME-POWER LATTICE DESIGNS

V. The analysis for a 6 × 6 incomplete lattice square illustrated with an example.*

W. T. FEDERER

Cornell University

I. INTRODUCTION

COCHRAN (1) has discussed the design and analysis for a number of incomplete lattice squares, that is, certain $k \times k$ lattice squares with less than $(k + 1)/2$ replicates. More recently, Kempthorne and Federer (2) have given the generalized analysis for $p \times p$, p a prime number, lattice squares with any number of replicates; also, they presented the design and analysis for a $k \times k$, k any integer, incomplete lattice square in multiples of 3 replicates. The incomplete lattice square would be suitable for $k^2 = 6^2$ or 10^2 varieties (or treatments) as well as for $k = 4, 5, 7, 8, 9$, etc. For the latter group a more efficient system of confounding may be found, and the design discussed, then, is suitable only for $k = 6$ or 10 and in some instances $k = 12$.

The purpose of this paper is to present the design and analysis for a $k \times k$ incomplete lattice square illustrated with a 6 × 6 incomplete lattice square in 3 and in 6 replicates.

II. DESCRIPTION OF THE NUMERICAL EXAMPLE

Uniformity trial yield data on corn (3) have been given in terms of pounds of ear corn for a 4 × 5 hill plot. From these data 6 replicates were constructed in such a way that the replicate was 6 plots wide and 6 plots long or 24 × 30 hills (Table 1). Since these are uniformity trial data, the so-called varieties are all the same but are treated as different entities in order to illustrate the numerical procedure.

*Contribution of the Biometrics Unit, Department of Plant Breeding, Cornell University, Plant Breeding paper No. 251.

In making up the design the variety numbers, 000, 015, $\cdots$, 541, 550, were assigned to the 6 rows and the 6 columns of plots in such a way that like numbers of the first digit, say i, were put in the same column and like numbers of the second digit, say j, were put in the same row in replicate I; like numbers of the second digit in the variety number, j, were included in the same column, and of the third digit, say h, were included in the same row in replicate II; and like numbers of the third digit, h, were put in the same column and of the first digit, i, were put in the same row in replicate III. After these arrangements had been made, the rows were randomized and then the columns were randomized in each replicate, thus retaining the above arrangement. For the 6 replicates in Table 1, replicate I and VI have the same arrangement of varieties in rows and columns; this is also true for replicates II and IV and for III and V. The rows and columns in the duplicate set of the 3 arrangements were subjected to a new randomization. The field arrangement of the 6^2 varieties and their yields in pounds of ear corn per 4×5 hill plot in each of the 6 replicates are presented in Table 1. The variety numbers are given in parentheses. The row, column, and replicate totals are given also. In practice, Table 1 need not be constructed if the above totals are obtained directly from the field books.

The arrangements in replicates I and VI, II and IV, and III and V are designated as the *ab*, *bc*, and *ca* arrangements, respectively. The above set of variety numbers is different from a pseudo-factorial arrangement in that of the $k^3 = 6^3$ possible combinations of 3 factors, a, b, and c, each at 6 levels, there are $k^2(k - 1) = 180$ missing combinations in the particular set of $6^2 = k^2$ variety numbers. The above method of numbering lends itself to punched card calculations and facilitates the design.

III. PROCEDURE OF COMPUTATIONS FOR REPLICATES I, II, AND III

A. *Analysis of Variance*

The analysis for a 6×6 incomplete lattice square in the *ab*, *bc*, and *ca* arrangements makes use of the row and column totals for each replicate and of the row and column totals of the variety totals in each of the 3 arrangements. The 3 systematic arrangements of variety numbers are given in Table 2 along with the totals of varieties from replicates I, II, and III. The row and column totals are given under the headings V_A, V_B, and V_C. The V_A totals represent the sum of $6 = k$ variety totals for which the first digit of the variety number remains constant. Likewise for the totals V_B and V_C. The sum of

TABLE I

FIELD ARRANGEMENT OF 6^2 VARIETIES IN AN INCOMPLETE LATTICE SQUARE WITH SIX REPLICATES*

Replicate I						Row total
(211) 33.9	(413) 32.1	(015) 33.1	(514) 29.2	(110) 31.1	(312) 29.6	189.0
(202) 28.7	(404) 31.2	(000) 34.4	(505) 32.3	(101) 30.8	(303) 31.1	188.5
(235) 30.4	(431) 30.5	(033) 29.7	(532) 28.1	(134) 31.3	(330) 30.4	180.4
(220) 31.1	(422) 33.0	(024) 29.2	(523) 28.8	(125) 30.7	(321) 29.3	182.1
(253) 32.0	(455) 39.4	(051) 32.4	(550) 24.2	(152) 31.3	(354) 30.9	190.2
(244) 31.1	(440) 30.0	(042) 28.5	(541) 29.4	(143) 26.1	(345) 28.3	173.4
187.2	196.2	187.3	172.0	181.3	179.6	1103.6

Replicate II						Row total
(541) 30.5	(321) 30.4	(211) 28.5	(051) 26.4	(101) 29.2	(431) 26.9	171.9
(143) 28.4	(523) 28.8	(413) 26.5	(253) 25.5	(303) 26.9	(033) 24.5	160.6
(042) 31.8	(422) 30.9	(312) 27.4	(152) 25.9	(202) 23.0	(532) 24.8	163.8
(244) 31.4	(024) 29.4	(514) 28.2	(354) 26.8	(404) 25.1	(134) 20.7	161.6
(440) 29.7	(220) 31.2	(110) 27.6	(550) 27.9	(000) 27.1	(330) 25.9	169.4
(345) 34.0	(125) 32.5	(015) 29.7	(455) 28.9	(505) 30.6	(235) 26.2	181.9
185.8	183.2	167.9	161.4	161.9	149.0	1009.2

Column total

Replicate III						Row total
(211) 30.3	(253) 28.1	(220) 31.1	(244) 27.1	(202) 27.2	(235) 25.8	169.6
(431) 31.2	(413) 30.4	(440) 30.0	(404) 27.9	(422) 26.6	(455) 24.3	170.4
(541) 30.3	(523) 30.8	(550) 30.8	(514) 27.1	(532) 27.0	(505) 24.6	170.6
(051) 29.2	(033) 30.9	(000) 31.4	(024) 28.9	(042) 28.6	(015) 26.3	175.3
(321) 27.9	(303) 28.6	(330) 30.0	(354) 28.1	(312) 30.9	(345) 31.1	176.6
(101) 31.6	(143) 29.9	(110) 30.6	(134) 29.0	(152) 32.5	(125) 30.4	184.0
180.5	178.7	183.9	168.1	172.8	162.5	1046.5

Replicate IV						Row total
(431) 27.0	(211) 24.6	(051) 27.5	(541) 29.0	(321) 24.5	(101) 27.3	159.9
(532) 28.8	(312) 26.1	(152) 27.4	(042) 28.3	(422) 27.9	(202) 27.5	166.0
(134) 27.6	(514) 27.8	(354) 26.9	(244) 26.4	(024) 25.4	(404) 28.1	162.2
(235) 27.9	(015) 27.7	(455) 30.9	(345) 27.9	(125) 24.5	(505) 26.1	165.0
(330) 25.6	(110) 26.7	(550) 28.3	(440) 26.4	(220) 25.4	(000) 25.1	157.5
(033) 28.5	(413) 27.8	(253) 26.4	(143) 23.1	(523) 28.7	(303) 26.5	161.0
165.4	160.7	167.4	161.1	156.4	160.6	971.6

Column total

Replicate V						Row total
(523) 27.0	(550) 26.1	(505) 24.8	(514) 23.8	(532) 23.8	(541) 24.7	150.2
(253) 29.1	(220) 28.4	(235) 24.4	(244) 22.4	(202) 24.4	(211) 27.3	156.0
(303) 30.2	(330) 29.0	(345) 30.5	(354) 23.1	(312) 24.8	(321) 25.6	163.2
(413) 32.3	(440) 28.0	(455) 28.9	(404) 28.8	(422) 27.4	(431) 28.3	173.7
(143) 31.0	(110) 33.5	(125) 28.9	(134) 27.1	(152) 27.5	(101) 30.1	178.1
(033) 32.0	(000) 31.2	(015) 31.7	(024) 31.0	(042) 29.8	(051) 28.7	184.4
181.6	176.2	169.2	156.2	157.7	164.7	1005.6

Replicate VI						Row total
(541) 28.7	(440) 30.8	(143) 29.3	(244) 28.5	(345) 28.7	(042) 31.5	177.5
(523) 29.2	(422) 28.7	(125) 28.7	(220) 29.2	(321) 28.5	(024) 27.4	171.7
(550) 27.3	(455) 27.6	(152) 26.8	(253) 29.6	(354) 29.2	(051) 27.5	168.0
(505) 26.2	(404) 30.8	(101) 28.3	(202) 32.1	(303) 29.9	(000) 28.4	175.7
(514) 24.7	(413) 26.8	(110) 27.7	(211) 29.3	(312) 28.3	(015) 29.9	166.7
(532) 25.3	(431) 26.0	(134) 27.6	(235) 29.4	(330) 28.0	(033) 28.2	164.5
161.4	170.7	168.4	178.1	172.6	172.9	1024.1

Column total

*Numbers in parentheses refer to variety number.

the V_A, V_B, or V_C must equal the grand total 3159.3, since all $108 = 3k^2$ plot yields enter into these sums. In practice V_A, V_B, and V_C totals need be constructed only once but the table was set forward to illustrate the detailed analysis.

The systematic arrangement of the plot yields is given in Table 3; in practice there would be little need of recopying the original yields. The appropriate row and column totals from Table 1 are given under the headings R_B and C_A, R_C and C_B, and R_A and C_C for replicates I, II, and III, respectively. The totals V_A, V_B, and V_C were copied from Table 2. It is imperative that R_u, C_u, and V_u, where $u = A$, B, or C be paired with respect to the varieties contained in each.

The column headed $V_B - 3R_B$ is obtained by subtracting 3 times the sum of varieties in a row of a replicate from the sum of the same 6 variety totals. The first $V_B - 3R_B$ value is

$$521.7 - 3(188.5) = -43.8.$$

The values for $V_A - 3C_A$, $V_C - 3R_C$, $V_B - 3C_B$, $V_A - 3R_A$, and $V_C - 3C_C$ are obtained similarly.

The first value in the next column headed $V_B - 2R_B - C_B = d$ is calculated as follows:

$$d = 521.7 - 2(188.5) - 161.9 = -17.2.$$

$C_B = 161.9$ was used since these same varieties appeared in R_B and V_B. The computations for the $V_A - 2C_A - R_A$, $V_C - 2R_C - C_C$, $V_B - 2C_B - R_B$, $V_A - 2R_A - C_A$ and $V_C - 2C_C - R_C$ values are obtained similarly, for example, the third $V_C - 2C_C - R_C = f$ is

$$f = 515.8 - 2(172.8) - 163.8 = 6.4.$$

The next to the last column represents the difference between the sum of the varieties in the row and of the same varieties when they appeared in columns. The values $R_B - C_B$, $R_C - C_C$, and $R_A - C_A$ are all computed similarly. The third $R_A - C_A = e$ value is

$$e = R_A - C_A = 169.6 - 187.2 = -17.6.$$

The computational procedure for obtaining the d', e' and f' values is described later. These values represent the adjustments for the variety means.

As partial checks each of the sums,

(i) $V_B - 3R_B$, $V_C - 3R_C$, $V_A - 3R_A$,
(ii) $V_A - 3C_A$, $V_B - 3C_B$, $V_C - 3C_C$,

TABLE 2

VARIETY TOTALS IN THE 3 ARRANGEMENTS USED TO CONSTRUCT THE GROUPINGS FOR THE FIRST 3 REPLICATES (VARIETY NUMBERS IN PARENTHESES)

Replicate I (*ab* arrangement)												V_B = total
(000)	92.9	(101)	91.6	(202)	78.9	(303)	86.6	(404)	84.2	(505)	87.5	521.7
(015)	89.1	(110)	89.3	(211)	92.7	(312)	87.9	(413)	89.0	(514)	84.5	532.5
(024)	87.5	(125)	93.6	(220)	93.4	(321)	87.6	(422)	90.5	(523)	88.4	541.0
(033)	85.1	(134)	81.0	(235)	82.4	(330)	86.3	(431)	88.6	(532)	79.9	503.3
(042)	88.9	(143)	84.4	(244)	89.6	(345)	93.4	(440)	89.7	(541)	90.2	536.2
(051)	88.0	(152)	89.7	(253)	85.6	(354)	85.8	(455)	92.6	(550)	82.9	524.6
V_A = Total	531.5		529.6		522.6		527.6		534.6		513.4	3159.3
Replicate II (*bc* arrangement)												V_C = total
(000)	92.9	(110)	89.3	(220)	93.4	(330)	86.3	(440)	89.7	(550)	82.9	534.5
(101)	91.6	(211)	92.7	(321)	87.6	(431)	88.6	(541)	90.2	(051)	88.0	538.7
(202)	78.9	(312)	87.9	(422)	90.5	(532)	79.9	(042)	88.9	(152)	89.7	515.8
(303)	86.6	(413)	89.0	(523)	88.4	(033)	85.1	(143)	84.4	(253)	85.6	519.1
(404)	84.2	(514)	84.5	(024)	87.5	(134)	81.0	(244)	89.6	(354)	85.8	512.6
(505)	87.5	(015)	89.1	(125)	93.6	(235)	82.4	(345)	93.4	(455)	92.6	538.6
V_B = Total	521.7		532.5		541.0		503.3		536.2		524.6	3159.3
Replicate III (*ca* arrangement)												V_A = total
(000)	92.9	(051)	88.0	(042)	88.9	(033)	85.1	(024)	87.5	(015)	89.1	531.5
(110)	89.3	(101)	91.6	(152)	89.7	(143)	84.4	(134)	81.0	(125)	93.6	529.6
(220)	93.4	(211)	92.7	(202)	78.9	(253)	85.6	(244)	89.6	(235)	82.4	522.6
(330)	86.3	(321)	87.6	(312)	87.9	(303)	86.6	(354)	85.8	(345)	93.4	527.6
(440)	89.7	(431)	88.6	(422)	90.5	(413)	89.0	(404)	84.2	(455)	92.6	534.6
(550)	82.9	(541)	90.2	(532)	79.9	(523)	88.4	(514)	84.5	(505)	87.5	513.4
V_C = Total	534.5		538.7		515.8		519.1		512.6		538.6	3159.3

(iii) the d values,
(iv) the f values, and
(v) the e values,

should equal zero. Other cross checks are available and are evident from the computations in Table 3.

The sums of squares for the analysis of variance of a 6×6 incomplete lattice square in 3 replicates are outlined below:

1. Correction term:

$$CT = \frac{(\text{Grand total})^2}{\text{total no. of plots}} = \frac{3159.3^2}{108} = 92{,}418.3008.$$

2. Total sum of squares:

$$33.9^2 + 32.1^2 + \cdots + 28.0^2 + 28.2^2 - CT$$

$$= 93{,}181.1500 - 92{,}418.3008 = 762.8492.$$

3. Replicate sum of squares:

$$\frac{1103.6^2 + 1009.2^2 + 1046.5^2}{36} - CT$$

$$= 92{,}543.8847 - 92{,}418.3008 = 125.5839.$$

4. Variety (ignoring row and column effects) sum of squares:

$$\frac{92.6^2 + 91.6^2 + \cdots + 92.6^2 + 82.9^2}{3} - CT$$

$$= 92{,}588.7767 - 92{,}418.3008 = 170.4759.$$

5. Error sum of squares for randomized complete blocks experiment:

$$762.8492 - 125.5839 - 170.4759 = 466.7894,$$

with $2(35) = 70$ degrees of freedom.

For the incomplete lattice square analysis, 5 additional sums of squares are required. These are presented below in steps 6, 7, 8, 9, and 10. The sums of squares from steps 6, 7, and 10 or from steps 8, 9, and 10 add to the sum of squares in step 5. The randomized complete blocks error sum of squares is partitioned into 3 parts, (i) row eliminating variety but ignoring column effects, (ii) column eliminating both variety and row effects, and (iii) the residual after removing the variety, row, column, and replicate effects, or alternatively, (i) column elimi-

TABLE 3

SYSTEMATIC ARRANGEMENT OF 6^2 VARIETY YIELDS IN REPLICATES I, II, AND III, TOTALS NECESSARY IN THE COMPUTATION OF THE ANALYSIS OF VARIANCE, AND ADJUSTMENTS FOR VARIETY MEANS. (VARIETY NUMBERS IN PARENTHESES)

Replicate I (*ab* arrangement)

	Variety nos. and yields						Total in row $=R_B$	V_B	V_B-3R_B	$V_B-2R_B-C_B=d$	d'	$R_B-C_B=e$	e'
	(000)	(101)	(202)	(303)	(404)	(505)							
	34.4	30.8	28.7	31.1	31.2	32.3	188.5	521.7	− 43.8	−17.2	+.01892	+26.6	+.56312
	(015)	(110)	(211)	(312)	(413)	(514)							
	33.1	31.1	33.9	29.6	32.1	29.2	189.0	532.5	− 34.5	−13.4	+.01474	+21.1	+.44669
	(024)	(125)	(220)	(321)	(422)	(523)							
	29.2	30.7	31.1	29.3	33.0	28.8	182.1	541.0	− 5.3	− 6.4	+.00704	− 1.1	−.02329
	(033)	(134)	(235)	(330)	(431)	(532)							
	29.7	31.3	30.4	30.4	30.5	28.1	180.4	503.3	− 37.9	− 6.5	+.00715	+31.4	+.66474
	(042)	(143)	(244)	(345)	(440)	(541)							
	28.5	26.1	31.1	28.3	30.0	29.4	173.4	536.2	+ 16.0	+ 3.6	−.00396	−12.4	−.26251
	(051)	(152)	(253)	(354)	(455)	(550)							
	32.4	31.3	32.0	30.9	39.4	24.2	190.2	524.6	− 46.0	−17.2	+.01892	+28.8	+.60970
C_A	187.3	181.3	187.2	179.6	196.2	172.0	1103.6			−57.1		+94.4	
V_A	531.5	529.6	522.6	527.6	534.6	513.4		3159.3					
V_A-3C_A	−30.4	−14.3	−39.0	−11.2	−54.0	−2.6			−151.5				
$f=V_A-2C_A-R_A$	−18.4	−17.0	−21.4	− 8.2	−28.2	−1.2	−94.4						
f'	−.36929	−.34119	−.42950	−.16457	−.56597	−.02408							

Replicate II (*bc* arrangement)

	Variety nos. and yields						R_C	V_C	V_C-3R_C	$V_C-2R_C-C_C=d$	d'	$R_C-C_C=e$	e'
	(000)	(110)	(220)	(330)	(440)	(550)							
	27.1	27.6	31.2	25.9	29.7	27.9	169.4	534.5	+26.3	+11.8	−.01298	−14.5	−.30696
	(101)	(211)	(321)	(431)	(541)	(051)							
	29.2	28.5	30.4	26.9	30.5	26.4	171.9	538.7	+23.0	+14.4	−.01584	− 8.6	−.18206
	(202)	(312)	(422)	(532)	(042)	(152)							
	23.0	27.4	30.9	24.8	31.8	25.9	163.8	515.8	+24.4	+15.4	−.01694	− 9.0	−.19053
	(303)	(413)	(523)	(033)	(143)	(253)							
	26.9	26.5	28.8	24.5	28.4	25.5	160.6	519.1	+37.3	+19.2	−.02112	−18.1	−.38318

	26.9	26.5	28.8	24.5	28.4	25.5	160.6	519.1	+37.3	+19.2	−.02112	−18.1	−.38318

Replicate II (*bc* arrangement)—*Continued*

	Variety nos. and yields						R_C	V_C	V_C-3R_C	$V_C-2R_C-C_C=d$	d'	$R_C-C_C=e$	e'
	(404)	(514)	(024)	(134)	(244)	(354)							
	25.1	28.2	29.4	20.7	31.4	26.8	161.6	512.6	+27.8	+21.3	−.02347	− 6.5	−.13760
	(505)	(015)	(125)	(235)	(345)	(455)							
	30.6	29.7	32.5	26.2	34.0	28.9	181.9	538.6	− 7.1	+12.3	−.01353	+19.4	+.41070
C_B	161.9	167.9	183.2	149.0	185.8	161.4	1009.2			+94.4		−37.3	
V_B	521.7	532.5	541.0	503.3	536.2	524.6		3159.3					
V_B-3C_B	+36.0	+ 28.8	−8.6	+56.3	−21.2	+40.4			+131.7				
$f=V_B-2C_B-R_B$	+ 9.4	+ 7.7	−7.5	+24.9	−8.8	+11.6	37.3						
f'	+.18866	+.15454	−.15052	+.49974	−.17662	+.23281							

Replicate III (*ca* arrangement)

	Variety nos. and yields						R_A	V_A	V_A-3R_A	$V_A-2R_A-C_A=d$	d'	$R_A-C_A=e$	e'
	(000)	(051)	(042)	(033)	(024)	(015)							
	31.4	29.2	28.6	30.9	28.9	26.3	175.3	531.5	+ 5.6	− 6.4	+.00704	−12.0	−.25404
	(110)	(101)	(152)	(143)	(134)	(125)							
	30.6	31.6	32.5	29.9	29.0	30.4	184.0	529.6	− 22.4	−19.7	+.02167	+ 2.7	+.05716
	(220)	(211)	(202)	(253)	(244)	(235)							
	31.1	30.3	27.2	28.1	27.1	25.8	169.6	522.6	+ 13.8	− 3.8	+.00418	−17.6	−.37259
	(330)	(321)	(312)	(303)	(354)	(345)							
	30.0	27.9	30.9	28.6	28.1	31.1	176.6	527.6	− 2.2	− 5.2	+.00572	− 3.0	−.06351
	(440)	(431)	(422)	(413)	(404)	(455)							
	30.0	31.2	26.6	30.4	27.9	24.3	170.4	534.6	+ 23.4	− 2.4	+.00264	−25.8	−.54619
	(550)	(541)	(532)	(523)	(514)	(505)							
	30.8	30.3	27.0	30.8	27.1	24.6	170.6	513.4	+ 1.6	+ 0.2	−.00022	− 1.4	−.02964
C_C	183.9	180.5	172.8	178.7	168.1	162.5	1046.5			−37.3		−57.1	
V_C	534.5	538.7	515.8	519.1	512.6	538.6		3159.3					
V_C-3C_C	−17.2	−2.8	−2.6	−17.0	+8.3	+51.1			+ 19.8				
$f=V_C-2C_C-R_C$	−2.7	+5.8	+6.4	+1.1	+14.8	+31.7	57.1						
f'	−.05419	+.11641	+.12845	+.02208	+.29704	+.63622							

nating variety but ignoring row effects, (ii) row eliminating both variety and column effects, and (iii) the residual.

6. Sum of squares for rows eliminating variety but ignoring column effects:

The within replicate sums of squares of the quantities $V_u - 3R_u$, $u = A, B,$ or C, yields the sums of squares for rows eliminating variety but ignoring the column effects with $3(k-1) = 15$ degrees of freedom,

$$\frac{(-43.8)^2 + (-34.5)^2 + \cdots + (-46.0)^2}{6k = 36} - \frac{(-151.5)^2}{6k^2 = 216}$$

$$+ \frac{(26.3)^2 + (23.0)^2 + \cdots + (-7.1)^2}{6k = 36} - \frac{(131.7)^2}{6k^2 = 216}$$

$$+ \frac{(5.6)^2 + (-22.4)^2 + \cdots + (1.6)^2}{6k = 36} - \frac{(19.8)^2}{6k^2 = 216}$$

$$= \frac{6945.19}{36} - \frac{22{,}952.25}{216} + \frac{4030.59}{36} - \frac{17{,}344.89}{216} + \frac{1278.52}{36}$$

$$- \frac{392.04}{216} = 152.0213.$$

7. Sum of squares for columns eliminating variety and row effects:

The sums of squares of the quantities $V_u - 2C_u - R_u$ within replicates yields the sum of squares for the column effects freed of variety and row effects,

$$\frac{(-18.4)^2 + \cdots + (-1.2)^2 + (-9.4)^2 + \cdots}{2k = 12}$$

$$\frac{+ (11.6)^2 + (-2.7)^2 + \cdots + (31.7)^2}{2k = 12}$$

$$- \frac{(-94.4)^2 + (37.3)^2 + (57.1)^2}{2k^2 = 72} = 169.3224,$$

with $3(k-1) = 15$ degrees of freedom.

8. Sum of squares for columns eliminating variety but ignoring row effects:

The within replicate sums of squares of the quantities $V_u - 3C_u$,

$$\frac{(-30.4)^2 + \cdots + (-2.6)^2 + (36.0)^2 + \cdots}{6k = 36}$$

$$\frac{+ (40.4)^2 + (-17.2)^2 + \cdots + (51.1)^2}{6k = 36}$$

$$- \frac{(-151.5)^2 + (131.7)^2 + (19.8)^2}{6k^2 = 216} = 267.9597,$$

yields the sum of squares for columns eliminating variety but ignoring row effects with $3(k - 1) = 15$ degrees of freedom.

9. Sum of squares for rows eliminating variety and column effects:

The within replicate sums of squares of the quantities $V_u - 2R_u - C_u$ represents the sum of squares attributable to row effects freed of varietal and column effects,

$$\frac{(-17.2)^2 + \cdots + (-17.2)^2 + (11.8)^2 + \cdots}{2k = 12}$$

$$\frac{+ (12.3)^2 + (-6.4)^2 + \cdots + (0.2)^2}{2k = 12}$$

$$- \frac{(-57.1)^2 + (94.4)^2 + (-37.3)^2}{2k^2 = 72} = 53.3842,$$

with $3(k - 1) = 15$ degrees of freedom.

10. Intrablock error sum of squares:

For the cases where $k \neq p$ a prime number the residual sum of squares representing the variation remaining after row, column, variety and replicate effects have been removed, must be computed by subtraction of the sums of squares obtained in steps 6 and 7 from that obtained in step 5, (or alternatively the difference between the sums of squares in step 5 and steps 8 and 9),

$$466.7894 - 152.0213 - 169.3224 = 145.4457$$

which equals, within rounding errors,

$$466.7894 - 267.9597 - 53.3842 = 145.4455,$$

with $2(k^2 - 1) - 6(k - 1) = 2(k - 1)(k - 2) = 40$ degrees of freedom.

The sums of squares are summarized in the analysis of variance table, Table 4. The mean squares are obtained in the usual manner.

B. *Estimation of Weights*

From the mean squares and the expectations of mean squares in Table 4, it is possible to obtain the various row, column, and intrablock estimated weights used for adjusting the variety means.

TABLE 4

ANALYSES OF VARIANCE FOR 3 REPLICATES

As a Randomized Complete Blocks Design

Source of variation	Degrees of freedom	Sum of squares	Mean square
Replicates	2	125.5839	62.7920
Varieties	35	170.4759	4.8707
Error	70	466.7894	6.6684
Total	107	762.8492	—

As an Incomplete Blocks Design

Source of variation	Degrees of freedom	Sum of squares	Mean square	d.f.	General case. Expectation of mean sq.
Replicates	2	125.5839	62.7920	2	—
Varieties (ignoring rows and columns)	35	170.4759	4.8707	$k^2 - 1$	—
Rows(elim.var. and ignoring columns)	15	152.0213	—	$3(k - 1)$	—
Columns (elim. var. and rows)	15	169.3224	11.2882	$3(k - 1)$	$\sigma_e^2 + \frac{1}{2}k\sigma_c^2$
Columns (elim. var. and ignoring rows)	15	267.9597	—	$3(k - 1)$	—
Rows (elim. var. and col.)	15	53.3842	3.5589	$3(k - 1)$	$\sigma_e^2 + \frac{1}{2}k\sigma_r^2$
Intrablock error	40	145.4457	3.6361	$2(k - 1)(k - 2)$	σ_e^2
Total	107	762.8492	—		

The reciprocal of the expected intrablock variance, σ_i^2 , is defined as the amount of information for intrablock comparisons and is estimated by the reciprocal of the intrablock error mean square,

$$w = \frac{1}{\hat{\sigma}_i^2} = \frac{1}{3.6361} = 0.27502.$$

Likewise, the reciprocals of the expected row $(\sigma_i^2 + k\sigma_r^2)$ and column $(\sigma_i^2 + k\sigma_c^2)$ variances are defined as the amount of information obtained on row, w_r , and column, w_c , effects, respectively.
They are estimated by

$$w_r = \frac{1}{2(\hat{\sigma}_i^2 + \frac{1}{2}k\hat{\sigma}_r^2) - \hat{\sigma}_i^2} = \frac{1}{3.4817} = 0.28722.$$

and

$$w_c = \frac{1}{2(\hat{\sigma}_i^2 + \frac{1}{2}k\hat{\sigma}_c^2) - \hat{\sigma}_i^2} = \frac{1}{18.9403} = 0.05280.$$

The amount of information obtained on row and column comparisons theoretically should be no greater than that on intrablock comparisons. Sample estimates of w_r or w_c are sometimes larger than w and whenever this is true the larger value should be taken equal to w. This will result in zero corrections to the mean for rows or columns or both. In this example w_r is larger than w and therefore should be taken equal to w, resulting in no adjustments for rows. *Since this is an illustrative example* the rule will *not* be followed and the estimates of w_r , w_c , and w as found will be used.

C. *Efficiency of the Incomplete Lattice Square Relative to Randomized Complete Blocks Design*

The average effective error variance per plot of this incomplete lattice square is

$$\left(\frac{3}{2}\right)\frac{2}{(k+1)}\left(\frac{3}{w_r + w_c + w} + \frac{(k-2)}{3w}\right)$$

$$= (3/7)(3(1.62593) + 4(1.21204)) = 4.16826.$$

The efficiency of this design relative to the randomized complete blocks design is the ratio of the average effective error variances, of the 2 designs,

$$\frac{6.6684}{4.16826} = 160 \text{ percent.}$$

Thus it appears that the 3 replicates in the incomplete lattice square gave as precise results as would have been obtained with about 5 replicates of a randomized complete blocks design. If the increase in efficiency had been small, say 0 to 10 or 15 percent, then the randomized complete blocks analysis should be used. In this case the adjustments should be made since an appreciable gain in efficiency resulted from use of the incomplete lattice square.

D. *Adjustments of Variety Means*

The adjusted variety means are obtained by adding 9 adjustments, 3 from each replicate, to the unadjusted variety means.* The adjustments, d', e', and f', are obtained by multiplying a factor times the d, e, and f values in Table 3, thus

$$d' = \frac{d(w - w_r)}{3k(w_r + w_c + w)},$$

$$e' = \frac{e(w_r - w_c)}{3k(w_r + w_c + w)},$$

and

$$f' = \frac{f(w - w_c)}{3k(w_r + w_c + w)}.$$

The first d', e', and f' values in replicate II of Table 3 were obtained as follows:

$$d' = +11.8(-.00110) = -.01298,$$

$$e' = -14.5(+.02117) = -.30696,$$

and

$$f' = +9.4(+.02007) = +.18866.$$

In the event that it is desired to adjust variety totals instead of means, multiply each of the factors by 3. As a partial check, the sum of the d', of the e', and of the f' values should equal zero.

The 9 adjustments for a variety mean are those in the same row

*O. Kempthorne has pointed out that the adjusted means may be obtained with 6 adjustments, 2 from each replicate. The adjustments in each replicate would be obtained by multiplying the $V_u - 3R_u$ values by $\frac{w - w_r}{3k(w_r + w_c + w)}$, and the $V_u - 3C_u$ values by $\frac{w - w_c}{3k(w_r + w_c + w)}$. These 6 adjustments are equivalent to the 9 presented here.

and column in each of the replicates as the variety plot yield. The adjusted mean for, say, variety 024, is

$$\frac{29.2 + 29.4 + 28.9}{3} + .00704 - .02329 - .36929 - .02343$$
$$- .13760 - .15052 + .00704 - .25404 + .29704$$
$$= 29.1667 - .64705 = 28.5196.$$

The remaining adjusted variety means in Table 5 were obtained similarly. The sum of the unadjusted means equals the sum of adjusted means within rounding errors.

TABLE 5

Sum of 9 adjustments	Adjusted mean	Sum of 9 adjustments	Adjusted mean
− .21972	30.7469	+ .16612	29.0328
+1.03307	30.7331	+ .31459	29.6146
− .64705	28.5196	− .47062	28.7294
+ .17312	28.5398	+ .57514	29.3418
−1.13840	28.4949	+ .36794	31.5012
+ .16365	29.4970	+ .77508	29.3751
+ .42685	30.9602	− .20281	27.8639
− .02052	29.7461	− .87577	28.7909
+ .60426	31.8043	−1.35531	28.8114
+1.04529	28.0453	− .01938	29.5139
−1.08767	27.0457	−1.92674	27.9733
+ .52005	30.4200	+ .78530	31.6520
− .10623	26.1938	+1.75015	30.9168
− .26343	30.6366	+ .69804	28.8647
−1.33881	29.7945	− .60293	28.8638
+1.40712	28.8738	+1.03867	27.6720
−1.10499	28.7617	− .57852	29.4882
− .31870	28.2146	+ .43336	28.0667
		+ .00005	1053.1003

E. *Standard Errors of a Mean Difference Between 2 Adjusted Variety Means*

The comparison of any 2 adjusted means of varieties which appear together in a row or column, such as 000 and 024, has a standard error of a mean difference of

$$\sqrt{\frac{2}{k}\left(\frac{2}{w_r + w_c + w} + \frac{k-2}{3w}\right)} = 1.643.$$

TABLE 6

VARIETY TOTALS IN THE 3 ARRANGEMENTS USED TO CONSTRUCT THE 3 GROUPINGS IN REPLICATES I TO VI FOR 6^2 VARIETIES IN AN INCOMPLETE LATTICE SQUARE (VARIETY NUMBERS IN PARENTHESES)

Replicates I and VI (*ab* arrangement)												V_B = Total
(000)	177.6	(101)	177.3	(202)	162.9	(303)	173.2	(404)	171.9	(505)	164.6	1027.5
(015)	178.4	(110)	177.2	(211)	173.9	(312)	167.1	(413)	175.9	(514)	160.8	1033.3
(024)	171.3	(125)	175.7	(220)	176.4	(321)	166.2	(422)	174.5	(523)	173.3	1037.4
(033)	173.8	(134)	163.3	(235)	164.1	(330)	168.9	(431)	169.9	(532)	157.8	997.8
(042)	178.5	(143)	167.8	(244)	166.9	(345)	180.5	(440)	174.9	(541)	172.6	1041.2
(051)	171.7	(152)	171.4	(253)	170.7	(354)	165.0	(455)	180.0	(550)	164.6	1023.4
V_A = Total	1051.3		1032.7		1014.9		1020.9		1047.1		993.7	6160.6
Replicates II and IV (*bc* arrangement)												V_C = Total
(000)	177.6	(110)	177.2	(220)	176.4	(330)	168.9	(440)	174.9	(550)	164.6	1039.6
(101)	177.3	(211)	173.9	(321)	166.2	(431)	169.9	(541)	172.6	(051)	171.7	1031.6
(202)	162.9	(312)	167.1	(422)	174.5	(532)	157.8	(042)	178.5	(152)	171.4	1012.2
(303)	173.2	(413)	175.9	(523)	173.3	(033)	173.8	(143)	167.8	(253)	170.7	1034.7
(404)	171.9	(514)	160.8	(024)	171.3	(134)	163.3	(244)	166.9	(354)	165.0	999.2
(505)	164.6	(015)	178.4	(125)	175.7	(235)	164.1	(345)	180.5	(455)	180.0	1043.3
V_B = Total	1027.5		1033.3		1037.4		997.8		1041.2		1023.4	6160.6
Replicates III and V (*ca* arrangement)												V_A = Total
(000)	177.6	(051)	171.7	(042)	178.5	(033)	173.8	(024)	171.3	(015)	178.4	1051.3
(110)	177.2	(101)	177.3	(152)	171.4	(143)	167.8	(134)	163.3	(125)	175.7	1032.7
(220)	176.4	(211)	173.9	(202)	162.9	(253)	170.7	(244)	166.9	(235)	164.1	1014.9
(330)	168.9	(321)	166.2	(312)	167.1	(303)	173.2	(354)	165.0	(345)	180.5	1020.9
(440)	174.9	(431)	169.9	(422)	174.5	(413)	175.9	(404)	171.9	(455)	180.0	1047.1
(550)	164.6	(541)	172.6	(532)	157.8	(523)	173.3	(514)	160.8	(505)	164.6	993.7
V_C = Total	1039.6		1031.6		1012.2		1034.7		999.2		1043.3	6160.6

The comparison of 2 adjusted means of varieties which do not appear together in a row or column has a standard error of a mean difference of

$$\sqrt{\frac{2}{k}\left(\frac{3}{w_r + w_c + w} + \frac{k-3}{3w}\right)} = 1.685.$$

The average standard error of a difference between any 2 adjusted variety means is

$$\sqrt{\frac{2}{k+1}\left(\frac{3}{w_r + w_c + w} + \frac{k-2}{3w}\right)} = 1.667.$$

Since the differences among the 3 standard errors are small, the average standard error of a mean difference may be used for all comparisons. A partial check is that the average standard error of a mean difference should have a value in between the other 2.

F. *Coefficient of Variation*

The coefficient of variation is the square root of the average effective error variance divided by the mean of the experiment,

$$\frac{\sqrt{4.16826}}{3159.3/108} = 7.0 \text{ percent.}$$

IV. PROCEDURE OF COMPUTATIONS FOR REPLICATES I TO VI

A. *Totals Required for the Analysis of Variance*

As explained previously the arrangement of varieties in replicates I and VI are identical except that the randomization of rows and columns differ. The grouping of varieties in the 2 replicates, I and VI, was denoted as the *ab* arrangement. Likewise, replicates II and IV and III and V are in the *bc* and *ca* arrangements, respectively. The totals of the 6 plot yields for each variety are presented in Table 6. The variety totals are arranged in *ab*, *bc*, and *ca* groupings in order to obtain the totals V_A , V_B , and V_C which are identical to the totals obtained in Table 2 except that each total in Table 6 comes from 36 individual items since there are 6 instead of 3 replicates. As before the totals V_A , V_B , and V_C need not be obtained twice in Table 6.

The next step is to construct a table (Table 7) similar to Table 3, the sole difference being that the total yield from the two plots of a variety are listed rather than the individual plot yields.

The remaining computations, $V_u - 3C_u$, $V_u - 3R_u$, $f = V_u - 2C_u - R_u$, $d = V_u - 2R_u - C_u$, and $e = R_u - C_u$, where $u = A$,

B, or C, are obtained as before. Also, the samc partial checks of computations for Table 3 are available for Table 6.

B. *Analysis of Variance*

As evident from the following steps the sums of squares in the analysis of variance are computed similarly to those for 3 replicates, the difference being that the number of replicates is now 6 instead of 3; the divisors for the sums of squares must be changed accordingly.

11. Correction term:

$$CT = \frac{(\text{Grand total})^2}{\text{total no. of plots}} = \frac{(6160.6)^2}{216} = 175{,}708.2980.$$

12. Total sum of squares:

$$33.9^2 + 32.1^2 + \cdots + 28.0^2 + 28.2^2 - CT$$

$$= 177{,}091.12 - 175{,}708.2980 = 1{,}382.8220.$$

13. Variety sum of squares:

$$\frac{177.6^2 + 177.3^2 + \cdots + 164.6^2}{6} - CT$$

$$= 175{,}904.3467 - 175{,}708.2980 = 196.0487.$$

14. Randomized complete block error sum of squares:

$$1382.8220 - 280.5514 - 196.0487 = 906.2219,$$

with $5(35) = 175$ degrees of freedom.

15. Sum of squares for rows eliminating variety but ignoring column effects:

$$\frac{(-65.1)^2 + \cdots + (-51.2)^2 + \cdots + (31.3)^2}{12k = 72}$$

$$- \frac{(-222.5)^2 + (218.2)^2 + (4.3)^2}{12k^2 = 432} = 154.8043,$$

with $3(k - 1) = 15$ degrees of freedom.

16. Sum of squares for columns eliminating both variety and row effects:

$$\frac{(-28.8)^2 + (-28.8)^2 + \cdots + (26.8)^2 + (33.0)^2}{4k = 24}$$

$$- \frac{(-146.9)^2 + (71.3)^2 + (75.6)^2}{4k^2 = 144} = 141.0065,$$

with $3(k - 1) = 15$ degrees of freedom.

17. Sum of squares for columns eliminating variety but ignoring row effects:

$$\frac{(-29.3)^2 + (-16.4)^2 + \cdots + (26.3)^2 + (48.2)^2}{12k = 72}$$

$$- \frac{(-222.5)^2 + (218.2)^2 + (4.3)^2}{12k^2 = 432} = 187.3860,$$

with $3(k - 1) = 15$ degrees of freedom.

18. Sum of squares for rows eliminating both variety and column effects:

$$\frac{(-23.4)^2 + (-6.7)^2 + \cdots + (-8.0)^2 + (18.7)^2}{4k = 24}$$

$$- \frac{(-75.6)^2 + (146.9)^2 + (71.3)^2}{4k^2 = 144} = 108.4248$$

with $3(k - 1) = 15$ degrees of freedom.

As a partial check, the sums of squares from steps 15 and 16 should equal the sums of squares from steps 17 and 18, within rounding errors,

$$154.8043 + 141.0065 = 295.8108.$$

$$187.3860 + 108.4248 = 295.8108.$$

19. Interaction sum of squares for row totals and replicates:

The interaction sums of squares of the row totals with replicates I and VI, II and IV, and III and V yields a pooled sum of squares with $3(2 - 1)(k - 1) = 15$ degrees of freedom. The varietal effects do not enter into this sum of squares since the row totals contain the same varieties in replicates in each row total of replicate I and of VI and likewise for the other 2 arrangements.

The interaction sum of squares of replicates I and VI and row totals is obtained from the two-way table constructed from the totals in Tables 1 and 7:

TABLE 7

SYSTEMATIC ARRANGEMENT OF 6^2 VARIETY YIELDS IN REPLICATES I AND VI, II AND IV, AND III AND V, TOTALS NECESSARY IN THE COMPUTATION OF THE ANALYSIS OF VARIANCE, AND ADJUSTMENTS FOR VARIETY MEANS (VARIETY NUMBERS IN PARENTHESES)

Replicates I and VI (*ab* arrangement)

	Variety nos. and yields from 2 plots						R_B	V_B	V_B-3R_B	$V_B-2R_B-C_B=d$	d'	$R_B-C_B=e$	e'
	(000)	(101)	(202)	(303)	(404)	(505)							
	62.8	59.1	60.8	61.0	62.0	58.5	364.2	1027.5	− 65.1	−23.4	−.30631	+41.7	+.06755
	(015)	(110)	(211)	(312)	(413)	(514)							
	63.0	58.8	63.2	57.9	58.9	53.9	355.7	1033.3	− 33.8	− 6.7	−.08770	+27.1	+.04390
	(024)	(125)	(220)	(321)	(422)	(523)							
	56.6	59.4	60.3	57.8	61.7	58.0	353.8	1037.4	− 24.0	− 9.8	−.12828	+14.2	+.02300
	(033)	(134)	(235)	(330)	(431)	(532)							
	57.9	58.9	59.8	58.4	56.5	53.4	344.9	997.8	− 36.9	− 6.4	−.08378	+30.5	+.04941
	(042)	(143)	(244)	(345)	(440)	(541)							
	60.0	55.4	59.6	57.0	60.8	58.1	350.9	1041.2	− 11.5	− 7.5	−.09818	+ 4.0	+.00648
	(051)	(152)	(253)	(354)	(455)	(550)							
	59.9	58.1	61.6	60.1	67.0	51.5	358.2	1023.4	− 51.2	−21.8	−.28536	+29.4	+.04763
C_A	360.2	349.7	365.3	352.2	366.9	333.4	2127.7			−75.6		+146.9	
V_A	1051.3	1032.7	1014.9	1020.9	1047.1	993.7		6160.6					
V_A-3C_A	−29.3	−16.4	−81.0	−35.7	−53.6	−6.5			−222.5				
$f=V_A-2C_A-R_A$	−28.8	−28.8	−41.3	−23.3	−30.8	+6.1	−146.9						
f'	−.42365	−.42365	−.60752	−.34274	−.45307	+.08973							

Replicates II and IV (*bc* arrangement)

	Variety nos. and yields from 2 plots						R_C	V_C	V_C-3R_C	$V_C-2R_C-C_C=d$	d'	$R_C-C_C=e$	e'
	(000)	(110)	(220)	(330)	(440)	(550)							
	52.2	53.3	56.6	51.5	56.1	56.2	326.9	1039.6	+58.9	+25.7	+.33641	−33.2	−.05378
	(101)	(211)	(321)	(431)	(541)	(051)							
	56.5	53.1	54.9	53.9	59.5	53.9	331.8	1031.6	+36.2	+22.8	+.29845	−13.4	−.02171
	(202)	(312)	(422)	(532)	(042)	(152)							
	50.5	53.5	58.8	53.6	60.1	53.3	329.8	1012.2	+22.8	+22.1	+.28929	− 0.7	−.00113
	(303)	(413)	(523)	(033)	(143)	(253)							
	53.4	54.3	57.5	53.0	51.5	51.9	321.6	1034.7	+69.9	+31.2	+.40841	−38.7	−.06269

Replicates II and IV (*bc* arrangement)—*Continued*

	Variety nos. and yields from 2 plots						R_C	V_C	V_C-3R_C	$V_C-2R_C-C_C=d$	d'	$R_C-C_C=e$	e'
	(404)	(514)	(024)	(134)	(244)	(354)							
	53.2	57.0	54.8	48.3	57.8	53.7	323.8	999.2	+27.8	+27.3	+.35736	− 0.5	−.00081
	(505)	(015)	(125)	(235)	(345)	(455)							
	56.7	57.4	57.0	54.1	61.9	59.8	346.9	1043.3	+2.6	+17.8	+.23300	+15.2	+.02462
C_B	322.5	328.6	339.6	314.4	346.9	328.8	1980.8			146.9		−71.3	
V_B	1027.5	1033.3	1037.4	997.8	1041.2	1023.4		6160.6					
V_B-3C_B	+60.0	+47.5	+ 18.6	+54.6	+0.5	+37.0			+218.2				
$f=V_B-2C_B-R_B$	+18.3	+20.4	+ 4.4	+24.1	−3.5	+7.6	+71.3						
f'	+.26919	+.30008	+.06472	+.35451	−.05148	+.11180							

Replicates III and V (*ca* arrangement)

	Variety nos. and yields from 2 plots						R_A	V_A	V_A-3R_A	$V_A-2R_A-C_A=d$	d'	$R_A-C_A=e$	e''
	(000)	(051)	(042)	(033)	(024)	(015)							
	62.6	57.9	58.4	62.9	59.9	58.0	359.7	1051.3	− 27.8	−28.3	−.37045	− 0.5	−.00081
	(110)	(101)	(152)	(143)	(134)	(125)							
	64.1	61.7	60.0	60.9	56.1	59.3	362.1	1032.7	− 53.6	−41.2	−.53931	12.4	+.02009
	(220)	(211)	(202)	(253)	(244)	(235)							
	59.5	57.6	51.6	57.2	49.5	50.2	325.6	1014.9	+ 38.1	− 1.6	−.02094	−39.7	−.06431
	(330)	(321)	(312)	(303)	(354)	(345)							
	59.0	53.5	55.7	58.8	51.2	61.6	339.8	1020.9	+ 1.5	−10.9	−.14268	−12.4	−.02009
	(440)	(431)	(422)	(413)	(404)	(455)							
	58.0	59.5	54.0	62.7	56.7	53.2	344.1	1047.1	+ 14.8	− 8.0	−.10472	−22.8	−.03694
	(550)	(541)	(532)	(523)	(514)	(505)							
	56.9	55.0	50.8	57.8	50.9	49.4	320.8	993.7	+ 31.3	+18.7	+.24478	−12.6	−.02041
C_C	360.1	345.2	330.5	360.3	324.3	331.7	2052.1			−71.3		−75.6	
V_C	1039.6	1031.6	1012.2	1034.7	999.2	1043.3		6160.6					
V_C-3C_C	−40.7	−4.0	+20.7	−46.2	+26.3	+48.2			+ 4.3				
$f=V_C-2C_C-R_C$	−7.5	+9.4	+21.4	−7.5	+26.8	+33.0	+75.6						
f'	−.11032	+.13827	+.31479	−.11032	+.39423	+.48543							

Row number	Rep. I	Rep. VI	V_B (Table 7)
0	188.5	175.7	364.2
1	189.0	166.7	355.7
2	182.1	171.7	353.8
3	180.4	164.5	344.9
4	173.4	177.5	350.9
5	190.2	168.0	358.2
Rep. total	1103.6	1024.1	2127.7

$$\frac{188.5^2 + 175.7^2 + \cdots + 168.0^2}{k = 6} + \frac{(2127.7)^2}{2k^2 = 72}$$

$$- \frac{(364.2)^2 + \cdots + (358.2)^2}{2k = 12} - \frac{(1103.6)^2 + (1024.1)^2}{k^2 = 36} = 39.8646,$$

with 5 degrees of freedom.

The pooled interaction sum of squares is

$$39.8646 + 28.4127 + 52.5323 = 120.8096.$$

20. Interaction sum of squares for column totals and replicates:

In the same manner as in the previous step, the pooled sum of squares of the interaction sums of squares of column totals with replicates is

$$17.9012 + 120.9327 + 37.7540 = 176.5879,$$

with $3(2 - 1)(k - 1) = 15$ degrees of freedom.

In the event that 9 replicates are used the interaction sums of squares of row totals and replicates and column totals and replicates are computed in the same manner and the pooled sums of squares will each contain $3(3 - 1)(k - 1)$ degrees of freedom.

21. Intrablock error sum of squares:

Subtraction of the sums of squares of steps 15, 16, 19, and 20 from the sum of squares of step 14 results in the intrablock error sum of squares with $(5k - 7)(k - 1) = 115$ degrees of freedom.

The sums of squares are summarized in Table 8. The expectation of various mean squares are given also.

C. *Estimation of Weights*

The weights represent quantity of information obtained for row,

TABLE 8

ANALYSES OF VARIANCE FOR 6 REPLICATES

As a Randomized Complete Block Design

Source of variation	Degrees of freedom	Sum of Squares	Mean Square
Replicates	5	280.5514	56.1103
Varieties	35	196.0487	5.6014
Error	175	906.2219	5.1784
Total	215	1,382.8220	—

As an Incomplete Blocks Design

Source of variation	Degrees of freedom	Sum of Squares	Mean Square	d.f.	General case. Expectation of m.s.
Replicates	5	280.5514	56.1103	5	
Varieties (ignoring rows and cols.)	35	196.0487	5.6014	$k^2 - 1$	
Rows (elim. var. and ignoring cols.)	15	154.8043	—	$3(k-1)$	
Columns (elim. var. and rows)	15	141.0065	9.4004	$3(k-1)$	$\sigma_i^2 + \frac{1}{2}k\sigma_c^2$
Cols. (elim. var. and ignoring rows)	15	187.3860	—	$3(k-1)$	
Rows (elim. var. and columns)	15	108.4248	7.2283	$3(k-1)$	$\sigma_i^2 + \frac{1}{2}k\sigma_r^2$
Row Interaction	15	120.8096	8.0540	$3(k-1)$	$\sigma_i^2 + k\sigma_r^2$
Column Interaction	15	176.5879	11.7725	$3(k-1)$	$\sigma_i^2 + k\sigma_c^2$
Intrablock Error	115	313.0136	2.7219	$(5k-7)(k-1)$	σ_i^2
Total	215	1,382.8220	—		

column and intrablock comparisons. The computations proceed in the same manner as for 3 replicates except that account must be taken of the fact that 6 replicates were used

$$w = \frac{1}{\hat{\sigma}_i^2} = \frac{1}{2.7219} = .36739,$$

$$w_r = \frac{1}{\frac{4}{3}\left(\frac{3(k-1)[(\hat{\sigma}_i^2 + k\hat{\sigma}_r^2) + (\hat{\sigma}_i^2 + \frac{1}{2}k\hat{\sigma}_r^2)]}{6(k-1)}\right) - \frac{1}{3}\hat{\sigma}_i^2}$$

$$= \frac{1}{9.2809} \doteq .10775,$$

and

$$w_e = \frac{1}{\frac{4}{3}\left(\frac{3(k-1)[(\hat{\sigma}_i^2 + k\hat{\sigma}_e^2) + (\hat{\sigma}_i^2 + \frac{1}{2}k\hat{\sigma}_e^2)]}{6(k-1)}\right) - \frac{1}{3}\hat{\sigma}_i^2}$$

$$\doteq \frac{1}{13.2080} = .07571.$$

D. *Efficiency of Incomplete Lattice Square Relative to Randomized Complete Blocks Design*

The average effective error variance per plot is

$$\left(\frac{6}{2}\right)\frac{2}{(k+1)}\left(\frac{3}{2(w_r + w_e + w)} + \frac{(k-2)}{6w}\right)$$

$$= (6/7)(3(0.90769) + 4(0.45365)) = 3.88943.$$

The efficiency relative to a randomized complete blocks design is

$$\frac{5.1784}{3.88943} = 133 \text{ percent.}$$

E. *Adjustment of Variety Means*

As before the values d, e, and f are multiplied by a factor to obtain the 9 adjustments for any variety mean,

$$d' = \frac{d(w - w_r)}{6k(w_r + w_e + w)} = .01309d,$$

$$e' = \frac{e(w_r - w_c)}{6k(w_r + w_c + w)} = .00162e,$$

and

$$f' = \frac{f(w - w_c)}{6k(w_r + w_c + w)} = .01471f.$$

For adjusting variety totals the factor 6(= number of replicates) in the denominator is omitted. The values d', e', and f' are given in Table 7.

The adjusted mean for variety 541 is the unadjusted mean plus the sum of the 9 adjustments associated with this mean,

$$\frac{29.4 + 30.5 + 30.3 + 29.0 + 24.7 + 28.7}{6} - .09818 + .00648$$

$$+ .08973 + .29845 - .02171 - .05148 + .24478 - .02041$$

$$+ .13827 = 28.7667 + .58593 = 29.3526.$$

The remaining means in Table 9 are adjusted in the same manner.

TABLE 9

Sum of 9 adjustments	Adjusted mean	Sum of 9 adjustments	Adjusted mean
−.59217	29.0078	− .23968	28.6270
+.20442	29.9377	+ .35372	28.2037
−.08469	28.4653	− .13106	27.5689
−.23937	28.7273	− .01306	28.1369
−.33514	29.4149	+ .09436	30.1777
−.50583	28.1109	+ .11934	27.6193
−.49743	29.0526	+ .18648	28.8365
−.51428	29.0190	− .10305	29.2136
−.24038	29.0429	− .03234	29.0510
+.12805	27.3447	+ .14042	28.4571
−.85065	27.1160	− .56560	28.5844
−.46585	28.1008	+ .02239	30.0224
−.05939	27.0906	+1.08758	28.5209
−.02148	28.9618	+1.32116	28.1212
−.56102	28.8390	+ .50894	29.3922
+.37042	27.7204	+1.23719	27.5372
−.08517	27.7315	+ .58593	29.3526
−.58330	27.8667	+ .36048	27.7938
		− .00006	1026.7663

As a partial check the sum of the unadjusted means should equal the sum of the adjusted means and the sum of the adjustments should add to zero within rounding errors.

F. *Standard Errors and the Coefficient of Variation*

The standard error of a mean difference between 2 varieties which appear in the same row and column is

$$\sqrt{\frac{2}{k}\left(\frac{2}{2(w_r + w_c + w)} + \frac{k-2}{6w}\right)} = 1.100.$$

The standard error of a mean difference between 2 varieties which do not appear together in the same row and column is

$$\sqrt{\frac{2}{k}\left(\frac{3}{2(w_r + w_c + w)} + \frac{k-3}{6w}\right)} = 1.167.$$

The average standard error of a mean difference between any 2 adjusted variety means is

$$\sqrt{\frac{2}{k+1}\left(\frac{3}{2(w_r + w_c + w)} + \frac{k-2}{6w}\right)} = 1.139.$$

The coefficient of variation is the square root of the average effective error mean square divided by the experimental mean,

$$\frac{\sqrt{3.88943}}{6160.6/216} = 6.9 \text{ percent.}$$

V. LITERATURE CITED

1. Cochran, W. G. Some additional lattice squares. Iowa Agricultural Experiment Station Research Bulletin 318, 731–747, 1943.
2. Kempthorne, O. and Federer, W. T. The general theory of prime-power lattice designs II. Designs for p^n varieties in p^s plots, and in squares. *Biometrics* 4, 109–121, 1948.
3. Zuber, M. S. A comparison of the relative efficiency of various experimental designs for corn yield tests. Unpublished M.S. Thesis, Iowa State College Library, Ames, Iowa. 1940.

AN APPLICATION OF SEQUENTIAL ANALYSIS TO WHITEFISH SAMPLING

G. B. OAKLAND

The University of Manitoba, Winnipeg

The Problem

THE WHITEFISH, Coregonus clupeaformis, sometimes contains cysts of the tapeworm, Triaenophorus crassus. The presence of these cysts in the edible parts of the fish is obnoxious but not harmful to man (1). When whitefish are exported from Canada it is desirable to know the infestation rate of the lot. This infestation rate is computed in the following manner: a certain number of fish are selected from the lot; they are examined for the number of cysts and the infestation rate is obtained from the formula:

$$\text{I.R.} = \frac{\text{total number of cysts in sample}}{\text{total weight of fish in sample in lbs.}} \times 100.$$

If the infestation rate is higher than a certain number, the lot is not suitable for export. If the infestation rate is lower than a certain number, the fish are exported. The test described above is a destructive test.

The infestation rates, calculated from the samples, vary as the carloads of fish are being sampled. Three of the factors which may cause this rate to change are:

1. The infestation rate varies from lake to lake and even in a large lake the infestation rate will vary from place to place.
2. The fish may not be selected at random when the sampling is being done. Hence as different individuals select the samples, bias may enter unless there is a random procedure in the sampling.
3. For lots of fish each having the same infestation rate, the infestation rates of the samples may vary greatly if the sample size is too small. Always present of course are the two types of errors: the probability of calling a good lot bad and the probability of calling a bad lot good. The probability of these two errors happening is decreased as the size of the sample is increased.

If the sampling plan has a fixed sample size, then when sampling lots whose infestation rates are not of borderline quality too many fish may be examined, resulting in a loss to the shipper. For lots which are of borderline quality, too few fish may be examined, resulting in a loss of confidence in the sampling plan.

A method of statistical inference which does not have a fixed sample size is sequential analysis (2), (3). In this, the fish are selected at random from the carload lots one at a time. The fish is cut open and filleted and the exposed parts are cut and examined for the presence of the cysts. The number of cysts is entered on a chart opposite the sample number (see figure 1). Whenever the plotted points go beyond the upper line the lot is rejected; if below the lower line, the lot is accepted and exported. As long as the plotted point is between the two lines the sampling continues.

With the sequential plan it is more feasible to work with cysts per fish rather than infestation rates. In commercial catches which are graded; round medium (1.5 and under 3 lbs.), round large (3 and under 4 lbs.), round jumbo (4 lbs. and over), the cysts per fish can be used in place of an infestation rate. For example, an infestation rate of 30 cysts per hundred lbs. of round large is equivalent to $(30 \times 3.5)/100 = 1.05$ cysts per fish, where 3.5 is the mid point of the range of the large classification. When dressed fish come through, a 15 per cent reduction in weight from the round figures has to be made. This would change the average weight for the large classification to 2.975 lbs. Then an infestation rate of 30 cysts per hundred lbs. of dressed large would be equivalent to $(30 \times 2.975)/100 = .8925$ cysts per fish.

For such a plan it is necessary to know the law of the distribution of cysts per fish. In the table below the results of the examination of 1188 medium whitefish in 1948 from a lake in western Canada are shown.

The terms of the negative binomial:

$1188(2.1440 - 1.1440)^{-.9308}$ fit the observed frequencies and a χ^2 of 2.82 is obtained with 6 degrees of freedom resulting in a probability greater than .71. This is the negative binomial, $(q - p)^{-k}$, which has been found useful in contagious studies (4) p. 54, and (5) and in fishing statistics has been found to fit very well the distribution of the number of cysts per fish. The negative binomial above has a mean, $kp = 1.0648$ and a variance, $kpq = 2.2830$.

The variance is computed by dividing the sum of squares of sample by $n - 1$, where n is size of sample, thus giving an estimate of the population variance. Then it is a simple matter to obtain

TABLE 1

Number of cysts	Number of fish	Number calculated by Negative Binomial
0	579	584
1	303	290
2	141	149
3	73	78
4	48	41
5	22	22
6	10	11
7	7	7
8	1	4
9	2	2
10	0	0
11	0	0
12	2	0
	1188	1188

$$q = \frac{kpq}{kp} = \frac{2.2830}{1.0648} = 2.1440.$$

since $q - p = 1$, hence $p = q - 1 = 1.1440$

and $$k = \frac{kp}{p} = \frac{1.0648}{1.1440} = .9308.$$

For the efficiency of this and other methods of estimating k, p, q, see p. 166 of (6).

The negative binomial as described by Wilks in (4) p. 54 is a discrete distribution where the variable x (the number of cysts) runs from 0 to infinity. The number of cysts can range in the same manner.

The Sequential Plan

The sequential plan as outlined by Wald (2) and applied to this problem is as follows. Let us assume we wish to classify lots coming through as being of 10 cysts per hundred lbs. or better (H_0) from those of 30 cysts per hundred lbs. or worse (H_1), assuming for purposes of convenient calculation that the average weight of the round fish is two lbs. and also that we are content to have the type I error, $\alpha = .05$ and the type II error, $\beta = .05$. To accommodate various sized shipments

which may vary from 1,200 to 30,000 lbs. α and β may be changed, that is, decrease them for larger shipments and increase them for smaller shipments. Under our two hypotheses H_0 and H_1 we then have two means,

H_0	H_1
$kp_0 = .2$	$kp_1 = .6$

From previous studies of the lakes in which the whitefish are caught one can estimate the variance of the cysts per fish. Here we shall assume that the variance in the population associated with H_1 is 2, i.e. $kp_1q_1 = 2$. In working out the sequential formulae below it was necessary to assume that k is the same under H_0 and H_1 . Anscombe (6) and Kendall (7) also found this helpful in connection with the negative binomial. In this case when k is kept constant and $kp_1q_1 = 2$ we are then able to obtain the values of the constants under the two hypotheses,

H_0	H_1
$kp_0 = .2$	$kp_1 = .6$
$kp_0q_0 = .3556$	$kp_1q_1 = 2$
$q_0 = 1 + p_0 = 1.7779$	$q_1 = 3.3333$
$p_0 = .7779$	$p_1 = 2.3333$
$k = .2571$	$k = .2571$

The variance under H_0 is lower than the variance under H_1 . There is some experimental evidence that as the mean number of cysts, kp, increases the variance, kpq, also increases.

Calculation of Acceptance and Rejection Lines

Mathematical calculations based upon Wald's discussion in (2) p. 126 yield the following formulae for the acceptance and rejection lines:

$$d = sn + h_0$$

$$\text{and } d = sn + h_1$$

where the slope of the lines, $$s = k \frac{\log \frac{q_1}{q_0}}{\log \frac{p_1 q_0}{p_0 q_1}} = .3439$$

and the intercepts, $$h_0 = \frac{\log B}{\log \frac{p_1 q_0}{p_0 q_1}} = -6.2659, \qquad \text{where } B = \frac{\beta}{1-\alpha}.$$

$$h_1 = \frac{\log A}{\log \frac{p_1 q_0}{p_0 q_1}} = 6.2659, \qquad \text{where } A = \frac{1-\beta}{\alpha}.$$

The derivation of the above formulae depends on the assumption that $p_1 > p_0$.

We accept H_0, if $d \leq .3439\ n - 6.2659$.

We accept H_1, if $d \geq .3439\ n + 6.2659$, where d is the cumulative number of cysts and n is the number of fish examined (see figure 1).

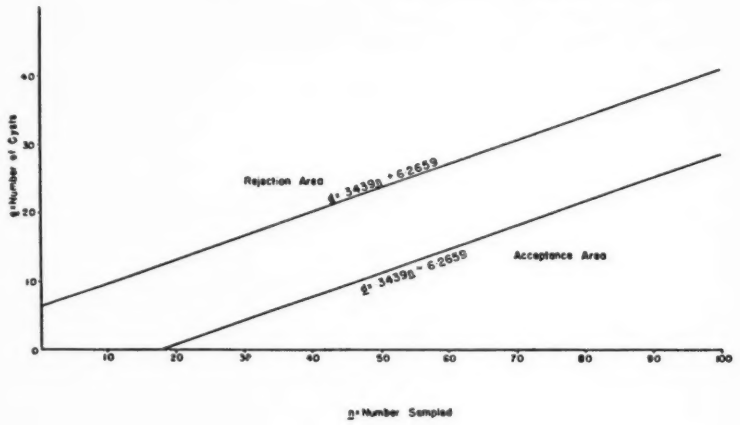

FIGURE 1

The operating characteristic curve

In order to know the probability $L(p)$ of accepting a lot of fish for any possible value of incoming quality, the operating characteristic curve is constructed. Here the probability, $L(p)$, is read from given

values of p (see figure 2). For example, when $p = .7779$ or the quality of fish coming through is at 10 cysts per hundred lbs., then the probability of accepting the lot is .95. When $p = 2.3333$ or a quality of 30 cysts per hundred lbs., the probability of accepting the lot is .05.

Mathematical calculations similar to those of (2) p. 162 give the operating characteristic curve for the negative binomial. The curve is described by:

$$L(p) = \frac{A^h - 1}{A^h - B^h}, \qquad \text{where} \qquad A = \frac{1-\beta}{\alpha},$$

$$\text{and} \qquad B = \frac{\beta}{1-\alpha},$$

$$\text{and} \qquad p = \frac{1 - \left(\frac{q_0}{q_1}\right)^h}{\left(\frac{p_1 q_0}{p_0 q_1}\right)^h - 1}, \qquad \text{where } h \text{ is a "dummy variable".}$$

For a given value of h, we compute the value of p and the value of $L(p)$ as shown in Table 2 and illustrated in figure 2.

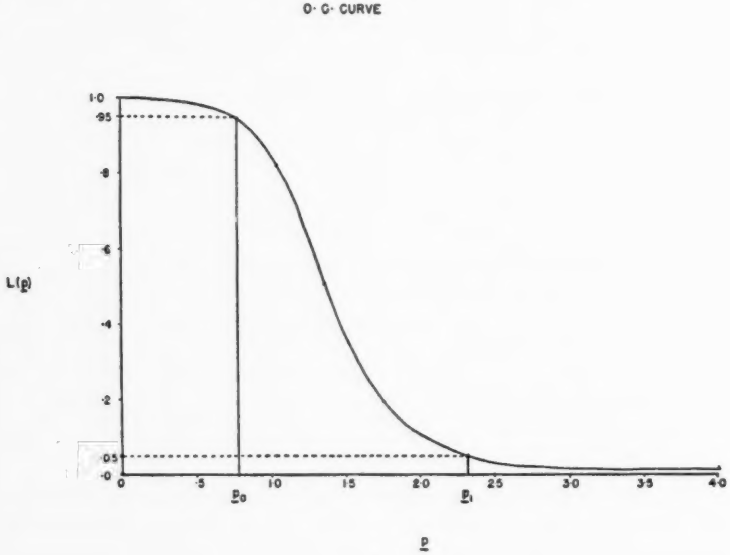

FIGURE 2

TABLE 2

h	p	$L(p)$
∞	0	1
1	.78	.95
1/2	1.02	.81
0	1.34	.50
−1/2	1.76	.19
−1	2.33	.05
$-\infty$	∞	.00

The Average Sample Number Curve

The average amount of inspection required for the test, $E(n)$, is given by the following formula, derived by calculations similar to those in (2) p. 164:

$$E(n) = \frac{h_1 + (h_0 - h_1)L(p)}{kp - s} .$$

Figure 3 which illustrates the Average Sample Number Curve (A.S.N.)

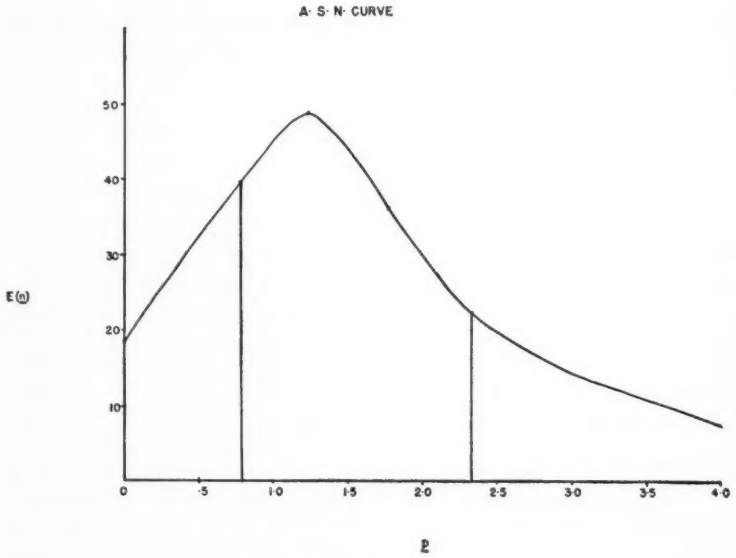

FIGURE 3

is based upon the following calculations in table 3:

TABLE 3

h	p	$E(n)$
∞	0	18.22
1	.78	39.19
1/2	1.02	48.61
−1/2	1.76	35.87
−1	2.33	22.02
−3/2	3.10	13.50
$-\infty$	∞	0

Under the above sequential scheme with lots coming in at 10 cysts per hundred lbs., the A.S.N. is 39 fish and at 30 cysts per hundred lbs., the A.S.N. is 22 fish. The maximum A.S.N. is 48 fish.

Advantages and Disadvantages

If the acceptance line equation:

$$d = .3439\, n - 6.2659$$

is solved for n by substituting $d = 0$, we find that if 16 fish were examined with no cysts present we would accept the hypothesis H_0.

If the rejection line equation:

$$d = .3439\, n + 6.2659$$

is solved for d when $n = 1$, we find that if 7 cysts or more are found in the first fish we accept the hypothesis H_1. This illustrates one of the advantages of the sequential plan. For fish of extremely good or extremely bad quality fewer samples have to be taken, but when fish is of borderline quality larger sample sizes are demanded by the scheme.

The sequential plan described above has one disadvantage under some working conditions. When the fish have to be cut in a location that may be a few blocks away from the carload lots of fish it is sometimes impossible to keep going back for more samples. In this case other sampling plans (single or double) may have to be used.

Mathematical Appendix

The derivation of the formula for p used in calculating the operating characteristic curve is given below.

The negative binomial distribution in which k has the same values under H_0 and H_1 is given by:

$$f(x,p) = q^{-k}\frac{\Gamma(x+k)}{x!\Gamma(k)}\left(\frac{p}{q}\right)^x \qquad \text{where} \qquad x = 0, 1, 2, \cdots$$

Then

$$\frac{f(x,p_1)}{f(x,p_0)} = \left(\frac{q_1}{q_0}\right)^{-k}\left(\frac{p_1q_0}{p_0q_1}\right)^x$$

The equation:

$$E\left[\frac{f(x,p_1)}{f(x,p_0)}\right]^h = 1$$

is solved for non zero h as on p. 137 of (2).

The equation involving the expected value becomes:

$$\sum_{x=0}^{\infty}\left[\left(\frac{q_1}{q_0}\right)^{-k}\left(\frac{p_1q_0}{p_0q_1}\right)^x\right]^h q^{-k}\frac{\Gamma(k+x)}{x!\Gamma(k)}\left(\frac{p}{q}\right)^x = 1$$

or

$$\left[\frac{1}{(p_0q_0)^h}\right]^{-k}\sum_{x=0}^{\infty}(q(p_0q_1)^h)^{-k}\frac{\Gamma(k+x)}{x!\Gamma(k)}\left(\frac{p(p_1q_0)^h}{q(p_0q_1)^h}\right)^x = 1$$

or

$$\left[\frac{1}{(p_0q_0)^h}\right]^{-k}[q(p_0q_1)^h - p(p_1q_0)^h]^{-k} = 1.$$

Raising both members of the above to the power $-1/k$ and simplifying we obtain on solving for p:

$$p = \frac{1-\left(\frac{q_0}{q_1}\right)^h}{\left(\frac{p_1q_0}{p_0q_1}\right)-1}.$$

REFERENCES

(1) Miller, R. B. The life history of the pike-whitefish tapeworm: triaenophorus crassus. Report prepared for Fisheries Branch, Department of Lands and Mines, Edmonton, March, 1945.

(2) Wald, A. Sequential tests of statistical hypotheses. *Annals of Mathematical Statistics*, Vol. XVI, No. 2, pp. 117–186, 1945.

(3) Statistical Research Group, Columbia University, *Sampling inspection: Principles, procedures and tables for single, double and sequential sampling in acceptance inspection and quality control.* McGraw-Hill Book Company, Inc., New York, 1947.

(4) Wilks, S. S. *Mathematical statistics.* Princeton University Press, Princeton, New Jersey, 1946.

(5) Fisher, R. A. The negative binomial distribution. *Annals of Eugenics* 11, 182, 1941.

(6) Anscombe, F. J. The statistical analysis of insect counts based on the negative binomial distribution. *Biometrics* 5, 2, pp. 165–173, 1949.

(7) Kendall, D. G. On some modes of population growth leading to R. A. Fisher's logarithmic series distribution. *Biometrika* 35, 6, 1948.

PROCEEDINGS OF THE SECOND INTERNATIONAL BIOMETRIC CONFERENCE

University of Geneva, Geneva, Switzerland

Tuesday 10:00 A.M., August 30, 1949 Formal Opening

The Conference was convened by Professor R. A. Fisher, President of the Biometric Society, who introduced Professor G. Tiercy, Rector of the University of Geneva. Professor Tiercy welcomed the Conference to the University and wished it every success. He was followed by Professor A. Franceschetti of the Faculty of Medicine at the University of Geneva. Speaking on behalf of the Swiss Academy of Medical Science and of the Local Arrangements Committee, Professor Franceschetti addressed the Conference as follows:

Monsieur le Recteur,

Messieurs les congressistes,

Mesdames,

Messieurs,

C'est pour moi un grand plaisir de vous souhaiter à tous la plus chaleureuse bienvenue.

Je tiens à exprimer nos remerciements les plus sincères au comité de la Société de Biométrie qui a bien voulu donner la préférence à notre ville comme siège de ce second congrès international de biométrie.

C'est pour nous un grand honneur de voir se rassembler à Genève les représentants le plus éminents de cette science, qui a le grand mérite d'attirer et de réunir tous ceux qui—dans les branches les plus diverses de la biologie, s'intéressent aux méthodes biométriques.

Il suffit de jeter un coup d'oeil au programme pour se rendre compte que la biométrie s'applique non seulement aux sciences biologiques et médicales, mais aussi à l'industrie et an particulier aux industries chimique et pharmacologique. Je suis persuadé que cette dernière, étroitement liée à la médecine, saura profiter encore davantage des méthodes biométriques, qu'elle a un peu trop négligées jusqu'à ces derniers temps. Un point présente pour toutes les branches un intérêt commun:

c'est la question de l'organisation des expériences, et de l'application à nos différentes disciplines des méthodes mathématiques et statistiques.

En tant qu'ophtalmologue s'intéressant tout particulièrement aux questions génétiques, qui jouent un rôle si important dans notre spécialité, j'aimerais souligner l'importance toujours grandissante de la biométrie génétique. Je ne rappellerai ici que les méthodes modernes permettant de calculer la fréquence des gènes, d'apprécier leur dominance relative, de rechercher le linkage, question si importante pour l'établissement futur de la carte chromosomique de l'homme.

Lorsque nous avons créé à la Clinique Ophtalmologique de Genève un service de génétique humaine—service que mon collaborateur, le Dr. Klein, montrera volontiers à tous ceux qui s'y intéressent—nous nous sommes rendu compte que nos travaux dans ce domaine exigeaient une collaboration plus étroite entre médecin et statisticien.

Je profite de l'occasion pour remercier ici notre Secrétaire du Congrès, qui a accepté l'hospitalité de la Clinique Ophtalmologique pour y installer son propre service et mettre ainsi à notre disposition toute sa compétence, dont nous profitons journellement. Cette collaboration entre médecin et statisticien demande évidemment une compréhension de la part de l'un et de l'autre. Elle exige de la part du statisticien une adaptation aux exigences spéciales de la biométrie génétique, et de la part du médecin, une préparation dans le domaine de la statistique mathématique, préparation qui malheureusement est encore insuffisante chez nous, et je crois dans bien d'autres pays aussi.

Je vous dans cette réunion internationale une possibilité d'attirer l'attention des autorités compétentes sur la nécessité de créer un enseignement mathématique et biologique, avant tout pour les futures biologistes et médecins.

Je souhaite que l'esprit de Genève ne manquera pas d'influencer favorablement cette collaboration internationale, ce qui assurera—j'en suis persuadé—un plein succès à ce second congrès international de biométrie.

Permettez-moi encore de remercier le Conseil d'Etat de la République et Canton de Genève, ainsi que le Conseil Administratif de la Ville de Genève qui, avec leur compréhension et leur bienveillance habituelles, nous ont conviés à une réception, jeudi prochain.

Nos remerciements vont également à Monsieur le Recteur pour ses aimables paroles de bienvenue, ainsi qu'aux donateurs généreux, qui de toutes parts en Suisse, nous ont aidés dans la réalisation pratique de ce congrès.

I would like to extend a hearty welcome to the eminent biometricians from all English speaking countries assembled here. We feel deeply honoured by the choice of Geneva as the meeting place for this second international conference of Biometrics. It may be said that in its origin biometry was essentially an Anglo-Saxon science. But research work in this branch is now rapidly spreading all over the world. This makes international cooperation and frequent exchanges of ideas a most important factor. I am sure that this conference will help to show the most efficient ways to apply biometrical methods.

As an ophthalmologist, especially interested in Genetics, a science which has taken a very important place in our specialty, I would like to emphasize that the question of teaching and education in Biometry is a most important one. We think that a thorough knowledge of Biometry is essential for all branches of biology and medicine.

I am sure that the application of Biometry to industrial work will become more and more important in the future.

The interesting papers announced on the program, the personal contacts that will be made and the exchange of ideas, will make this conference attractive for everybody.

We hope that the traditionally international spirit of Geneva will help in making this conference a success.

E'con vivo piacere che porgo a nome anche del Comitato d'organizzazione il saluto piu' cordiale ai partecipanti di questo secondo congresso internazionale di Biometria.

Siamo lieti di veder riuniti a Ginevra—città che si compiace di poter contribuire alla soluzione dei problemi internazionali pacifici e scientifici, creando quello spirito di collaborazione necessaria tanti rappresentanti illustri di questa scienza relativamente moderna che conta gia' un numero considerevole di scienziati in tutti i campi della biologia, della matematica e anche dell'industria.

Sono convito che de questo congresso, oltre che al successo scientifico assicurato gia' per il programma interessantissimo delle relazioni e comunicazioni annunciate, ne risulteranno dei contatti personali, con scambio di idee fra i rappresentati dei differenti paesi e delle diverse branche scientifiche, che contribuiranno certamente al sempre maggiore sviluppo delle nostre conoscenze in tutti i campi della biometria.

Following Professor Franceschetti's address, the Conference recessed briefly and then reconvened for its first business meeting with President Fisher in the chair. The Secretary reported on the present status of the Biometric Society—its membership, directory, quarters and international affiliations. The chairman asked G. Darmois, W. G. Cochran and G. Rasch to serve as a Committee on Resolutions and A. Linder, J. W. Hopkins and C. I. Bliss as a Finance Committee. The Conference next considered the standardization of statistical nomenclature and symbols. D. van Dantzig submitted a proposal which had been prepared at the Mathematisch Centrum in Amsterdam. Since other groups were reported to be working on the same problem, the question was referred to a special committee consisting of F. Yates (Chairman), G. M. Cox, D. van Dantzig, E. C. Fieller, R. A. Fisher and E. C. Wood. C. I. Bliss then summarized the report of a Sub-committee of the U. S. National Research Council which recommended that UNESCO establish an international mathematical and statistical consulting center provided with adequate computational equipment. He suggested that if the Conference favored such a center, it pass a resolution to this effect. Following extended discussion, the Conference voted to delay action on the proposal until the closing business meeting. After announcements on Conference procedure the meeting adjourned.

Friday 4:00 P.M., September 2, 1949 Business Meeting

The Conference closed with a second business meeting under the chairmanship of President Fisher. On behalf of his committee Professor Darmois proposed and the Conference adopted unanimously the following resolution:

"La commission des résolutions a, comme vous le savez, le fort agréable devoir de présenter, si j'ose dire, les fleurs de notre reconnaissance a tous ceux qui ont permis à ce congrès de se réunir, de travailler, et mème de s'amuser un peu.

"Tout d'abord nous remercions Mr. le Recteur Tiercy qui nous a, avec tant de

bonne grâce, donné un logis. Il avait des craintes légitimes, dues à une première et fâcheuse expérience. Espérons que les espoirs que lui inspirait la métrie n'ont pas été déçus, et que, si j'ose risquer un innocent calembour, notre métrie à nous aura été la mesure.

"Après avoir parlé du lieu de notre labeur, venons à ceux de nos délassements. Mr. le Professeur Franceschetti et Madame Franceschetti nous ont fait passer dans leur si belle propriété des instants charmants. Nous les remercions bien vivement.

"Nous remercions aussi profondément le Conseil d'Etat de la République et Canton de Genève et le Conseil Administratif de la Ville de Genève qui ont manifesté l'intérêt qu'ils portent à nos travaux en nous recevant au Palais Eynard, et nous gardons le meilleur souvenir de cette réunion si cordiale, où se trouvaient conviés également nos amis et collègues du Congrès de la Population.

"Nous pensons bien aussi qu'il existe de modestes et généreux donateurs qui ont permis l'organisation et la réussite de ce Congrès. Qu'ils trouvent ici l'expression de notre reconnaissance.

"Et je voudrais remercier, avec une chaleur toute particulière, l'infatigable organisateur de notre Réunion, celui grâce à qui elle fut un succès, Mr. le Professeur Linder, qui par un curieux phénomène d'ubiquité, réussit à se trouver, toujours aussi aimable qu'efficace, partout où l'on a besoin de lui.

"Je ne finirai pas sans dire toute notre reconnaissante sympathie à notre interprète, M. Walter Keiser, qui a si vaillament lutté dans un combat continu, et qui pouvait sembler inégal, avec se mélange insidieux et bilingue de statistique et de biologie. Je dois dire qu'à notre étonnement, il en est souvent sorti victorieux."

Dr. Yates reported for his committee the following resolution, which was adopted unanimously:

"The Biometric Society notes with interest that many groups are working on the problem of standardization of biometrical symbols. It believes that much more work of this kind will be required and expresses the hope that the International Statistical Institute will take the lead in the establishment of an internationally acceptable terminology and notation."

The Secretary then read a resolution on the proposed "International Computation Center", which was discussed, amended from the floor and adopted unanimously in the following form:

"Whereas it has been proposed to UNESCO by a U. S. National Research Council Subcommittee that an international mathematical and statistical consulting center be established by the United Nations under the sponsorship of UNESCO and other international bodies, to be staffed with theoretical and applied mathematicians and statisticians and provided with adequate computational equipment and aids, and

"Whereas the Biometric Society is an international society devoted to advancing the mathematical and statistical aspects of biology and is interested in the wise use of international funds for these purposes,

"Be it resolved that the Biometric Society strongly urge the appointment of an international committee by UNESCO, to examine the position and recommend the most effective means of furthering mathematical and statistical methods.

"Therefore, in the opinion of this Society, full attention should be given to the original proposal, it being understood that the statistical and biometrical section will require a separate organization of personnel and budget. If such a center is established, the Society is of the opinion that it would be best located in a country where,

owing to the lack of adequate statistical institutions, it would serve the needs of that region. Consideration should also be given to the use of international funds to develop existing organizations so that together they may serve more perfectly international ends.

"The Society is also of the opinion that the statistical organization of the United Nations and of its specialized agencies should be strengthened by the appointment of a professional staff qualified to further the application of adequate statistical techniques in the collection, analysis and critical evaluation of data.

"Immediate needs in the opinion of the Society are:

(a) Translations of specialized scientific books and articles requested by countries feeling this need and
(b) Studentships to graduate students anxious to pursue biometrical, biomathematical and statistical studies in countries more advanced in such fields than their own.

These proposals are particularly relevant to the field of biometrical studies owing to the rapid progress made in recent years in a few countries. The Society therefore strongly supports the proposals in these two directions put forward independently to UNESCO in the provisional report of the International Statistical Institute on the 'Furtherance of Statistical Education'."

Professor Felix Bernstein moved and the Conference passed the following resolution:

"The Conference recommends that every future student of science, education, medicine, administration and engineering, in private or public schools, should be brought into contact with the concepts of probability and statistics at the undergraduate level, and preferably in the Freshman or Sophomore years or their equivalent."

On behalf of ten members of the Conference, Professor Darmois submitted several recommendations for the next International Biometric Conference, among them (1) separate general and specialized sessions, (2) greater emphasis upon the application of statistics to concrete problems and (3) simultaneous translation of the proceedings. Their suggestions were filed for the use of the next Conference Committee. The Conference then adjourned.

Registration at the Conference

The final registration at the Conference totalled 103. Official delegates of national governments were A. S. Agapitedès, Greece; S. T. Bok and G. van Iterson, the Netherlands; F. Bosquet and Jean Landrain, Belgium; P. De Alba, Mexico; C. Gini and G. Pompilj, Italy; T. Kemp, Denmark; A. V. M. P. de Magalhães, A. M. de A. Pyrrait and A. M. D. Tovar de Lemos, Portugal; V. Montoya, Venezuela; S. Rios, Spain. Other delegates were R. T. H. Beverton and T. Dalling, Ministry of Agri-

culture and Fisheries, England; *A. Buzzati-Traverso, National Research Council, Italy; F. Chodat, International Union of Biological Sciences; *G. M. Cox, Eastern North American Region, Biometric Society; D. van Dantzig, Mathematisch Centrum, Amsterdam; Mme. E. Defrise, l'Institut National des Sciences Naturelles de Belgique; A. Fleisch, l'Académie Suisse des Sciences médicales; R. Gautier and M. Pascua, World Health Organization; P. Hartley, The Royal Society, London; *J. W. Hopkins, National Research Council, Canada; *A. Linder, International Statistical Institute; L. Martin, l'Institut agronomique de Gembloux; E. Morice, Ministère de la Santé Publique, Paris; *M. G. Neurdenburg, Municipality of Amsterdam.

Of the other participants, the following were members of the Biometric Society as of July 1949: H. Åstrand, M. S. Bartlett, E. Beccari, M. W. Bentzon, F. Bernstein, C. I. Bliss, E. Boeri, F. Brambilla, E. Broomfield, L. L. Cavalli, W. G. Cochran, G. Darmois, O. L. Davies, B. B. Day, P. J. Delaporte, C. Dieulefait, R. Feron, E. C. Fieller, D. J. Finney, R. A. Fisher, A. Franceschetti, A. Hald, J. B. S. Haldane, M. J. R. Healy, J. O. Irwin, N. K. Jerne, F. B. Leech, J. M. Legay, O. U. Maaløe, K. B. Madhava, P. C. Mahalanobis, C. A. G. Nass, W. L. M. Perry, L. Pirocchi, M. H. Quenouille, A. Rapoport, G. Rasch, E. C. Reeve, M. P. Schützenberger, D. Schwartz, R. Scossiroli, K. L. Smith, L. Tônrqvist, V. Tonolli, J. M. Tripod, A. Vessereau, J. M. Vincent, N. Visconti di Modrone and F. Yates. The list of participants was completed by R. Borth, G. Deloffre, E. Dieulefait, E. F. Drion, H. Florin, N. F. Gjeddebaek, G. Goudswaard, R. Jardine, Dr. Jenny, A. Jensen, R. Karschon, D. Klein, E. van der Laan, A. Lubin, Mrs. H. W. Norton P. S. Østergaard, J. Reboul, S. A. Rice, S. Rosin, S. N. Roy, H. J. Stutvoet, A. A. Weber and W. Wegmüller.

Thus the participants at the Conference included 14 delegates from the governments of 9 different countries, 16 delegates from scientific or governmental organizations and 73 other conferees. The registration also showed 53 members of the Biometric Society (as of July 1949). Nineteen different countries were represented by the participants as summarized in the last issue of BIOMETRICS.

---

*Members of the Biometric Society as of July 1949.

THE PRESENT STATUS OF BIOMETRY

WILLIAM G. COCHRAN

Johns Hopkins University

Abstract

In the founding of the Biometric Society, a wise decision was taken to regard biometry as roughly synonymous with quantitative biology and as a discipline requiring the cooperation of biologists, mathematicians, and statisticians.

It seems appropriate to present some review of the major current areas of research. The field has grown to such an extent, however, that this review must be based on a small and biased sample of the literature, a defect which it is hoped will be remedied in the discussion following the paper. For this review the field is divided into two sections: *statistical techniques*, which comprise the design of experiments, sampling techniques and methods for the analysis of data, and *biological applications*, which deal with the uses made of these tools in biology.

The principles of sound experimental design are gradually coming to be appreciated, though dissemination is still uneven. In recent years many papers have appeared which apply these principles to a wide range of experimentation, and serve a valuable expository purpose. Examples of the topics dealt with are microbiological assays (Wood and Finney, 1946), the effect of streptomycin on pulmonary tuberculosis (British Medical Research Council, 1948), remedies for headaches (Jellinek, 1946), the assay of tetanus antitoxin (Ipsen, 1942), the assay of penicillin (Bliss, 1946), feeding trials on milk cows (Cochran et al, 1941), tests of mosquito repellents (Wadley, 1946), and the stimulation of denervated muscle by electric currents (Solandt et al, 1943). The newer developments are mainly specialized devices for decreasing the work required in an experiment. Three techniques are described—the fractional replication of large factorial experiments (Finney, 1945, 1946), the "up and down" method for sensitivity experiments (Dixon and Mood, 1948) and sequential experimentation (Wald, 1947).

Sampling theory has advanced to the point where good sampling is relatively simple, provided that the population can be divided into accessible sampling-units and that satisfactory methods of measurement have been devised. Many problems remain where these conditions are

not met, as in sampling from bulked material that is heterogeneous and in obtaining information from human populations. A clearer understanding of the properties of systematic sampling (Yates, 1948 and Matérn, 1947) is now available. One special problem on which progress has been made is that of estimating the total of a mobile population (e.g. of insects) in some natural area, taking account of births, deaths, immigration and emigration (Jackson, 1948, Fisher and Ford, 1947).

Recent literature on the analysis of data is voluminous. For the most part it is concerned with the extension of well-established techniques rather than with basically novel approaches. As might be expected, techniques are becoming more complex. Problems still tend to be stated and solved in the terms that are most convenient to the statistician, rather than with the assumptions that are most realistic for biological use. Two lines of work are described. That on discriminant functions (Rao, 1948, Cochran and Bliss, 1948, C. A. B. Smith, 1947) illustrates the application of a relatively complex technique. Work on human genetics (Bernstein, 1931, Haldane, 1934, Fisher, 1934, 1935, 1940, Finney, 1940, 1948) has produced ingenious methods for estimating gene ratios and linkage intensities from the type of non-random sample that arises in this field.

The extent to which profitable use is made of statistical techniques in biology varies greatly from one branch to another. For instance, in plant breeding, the applied biologists appear to be alert in adopting new biometric methods, but in bioassay and in clinical medicine the penetration of statistical ideas is more sporadic and incomplete. Much experimentation evidences little appreciation of the scientific method. Where biological measurements are difficult, there is a tendency to overlook important sampling problems that are present.

Progress in the formulation of biological laws lags behind that in physics and chemistry. Possible reasons are presented. Recent methods of attack are illustrated by discussion of the discovery of the logarithmic series distribution (Fisher et al, 1943, Williams, 1947), the rhesus factor (Fisher, 1947), and the relation of weather to plant yields.

Mathematical biology, or in other words, the construction and investigation of mathematical models for biological processes, shows increased activity. New applications to neurophysiology are noteworthy (Householder and Landahl, 1944, Wiener and Rosenbluth, 1946, McCulloch and Pitts, 1948). As Wiener (1948) has pointed out, this work has applications in many other branches of science. Other topics treated are binocular vision (Luneberg, 1947), dark adaptation (Moon and Spencer, 1945) and the Weber-Fechner law (von Schelling, 1944), in addition to further work on epidemic theory (Wilson and Worcester,

1945) and on evolution (Malécot, 1946). Recent interest in stochastic processes may be beneficial in leading to more realistic mathematical models of biological processes. In the past most models have been deterministic even for phenomena that are obviously affected by chance.

To summarize, biometry is flourishing. Research on statistical techniques is ample in volume. To avoid the danger of losing touch with the biologist, there is need for investigation of simpler techniques based on assumptions that are reasonable for biological applications, and for work on a greater diversity of problems. The use of statistical methods by biologists is likely to increase: poor training in scientific method appears a greater obstacle than weakness in mathematics. The accomplishments in mathematical biology are not as well known or as influential as they might be. Increased effort might profitably be devoted to this field.

The Biometric Society has the opportunity to play an important role in the future development of biometry. M. Teissier's reminder that our society is an experiment in cooperation, and one where success will be difficult to attain, is wise. I hope that the scope of the society will continue to be conceived broadly and that biologists and mathematicians will have their share of responsibility in determining our future activities. Finally, I hope that we will set an example in international cooperation.

SELECTED REFERENCES

Bernstein, F. (1931), Zur Grundlegung der Chromosomentheorie der Vererbung beim Menschen mit besonderer Berücksichtigung der Blutgruppen. *Z. indukt. Abstamm. u. Vererbungslehre*, 57, pp. 113-138.

Bliss, C. I. (1946), A revised cylinder-plate assay for penicillin. *Jour. Amer. Phar. Assoc.*, 35, pp. 6–12.

British Medical Research Council (1948), Streptomycin treatment of pulmonary tuberculosis. *Brit. Med. Jour.*, 2, pp. 769–783.

Cochran, W. G., Cannon, C. Y. and Autrey, K. M. (1941), A double change-over design for dairy cattle feeding experiments. *Jour. Dairy Sci.*, 24, pp. 937–951.

—— and Bliss, C. I. (1948), Discriminant functions with covariance. *Ann. Math. Stat.*, 19, pp. 151–176.

Dixon, W. J. and Mood, A. M. (1948), A method for obtaining and testing sensitivity data. *Jour. Amer. Stat. Assoc.*, 43, pp. 109–126.

Finney, D. J. (1940), The detection of linkage. *Ann. Eugen.*, 10, pp. 171–214.

—— (1945), The fractional replication of factorial arrangements. *Ann. Eugen.*, 12, pp. 291–301.

—— (1946), Recent developments in the design of field experiments. III. Fractional replication. *Jour. Agr. Sci.*, 36, pp. 184–191.

—— (1948), The estimation of gene frequencies from family records. I. Factors without dominance. *Heredity*, 2, pp. 199–218.

Fisher, R. A. (1934), The effect of methods of ascertainment upon the estimation of frequencies. *Ann. Eugen.*, 6, pp. 13–25.

—— (1935), The detection of linkage with "dominant" abnormalities. *Ann. Eugen.*, 6, pp. 187–201.

—— (1940), The estimation of the proportion of recessives from tests carried out on a sample not wholly unrelated. *Ann. Eugen.*, 10, pp. 160–170.

——, Corbet, A. S. and Williams, C. B. (1943), The relation between the number of species and the number of individuals in a random sample of an animal population. *Jour. An. Ecol.*, 12, pp. 42–58.

—— (1947), The Rhesus factor. *Amer. Scientist*, 35, pp. 95–103.

—— and Ford, E. B. (1947), The spread of a gene in natural conditions in a colony of the moth panaxia dominula L. *Heredity*, 1, pp. 143–174.

Haldane, J. B. S. (1934), Methods for the detection of autosomal linkage in man. *Ann. Eugen.*, 6, pp. 26–65.

Householder, A. S. and Landahl, H. D. (1944), *Mathematical biophysics of the central nervous system.* Principia Press, Bloomington, Indiana.

Ipsen, J. (1942), Systematische und zufällige Fehlerquellen bei Messung kleiner Antitoxinmengen. *Zeitschrift für Immunitätsforschung*, 102, pp. 347–368.

Jackson, C. H. N. (1948), The analysis of a tsetse-fly population. III. *Ann. Eugen.*, 14, pp. 91–108.

Jellinek, E. M. (1946), Clinical tests on comparative effectiveness of analgesic drugs. *Biometrics*, 2, pp. 87–91.

Luneburg, R. K. (1947), *Mathematical analysis of binocular vision.* Princeton University Press.

Malécot, G. (1946), La consanguinité dans une population limitée. *C. R. Acad. Sci. Paris*, 222, pp. 841-843.

Matérn, B. (1947), Methods of estimating the accuracy of line and sample plot surveys. *Medd. fr. Statens Skagsforsknings Inst.*, 36.

McCulloch, W. S. and Pitts, W. (1948), The statistical organization of nervous activity. *Biometrics*, 4, pp. 91–99.

Moon, P. and Spencer, D. E. (1945), Visual dark adaptation: a mathematical Formulation. *Jour. Math. Phys. M. I. T.*, 24, pp. 65–105.

Rao, C. R. (1948), The utilization of multiple measurements in problems of biological classification. *Jour. Roy. Stat. Soc.* Ser. B., 10, pp. 159–203.

Smith, C. A. B. (1947), Some examples of discrimination. *Ann. Eugen.*, 13, pp. 272-282.

Solandt, D. Y., DeLury, D. B. and Hunter, J. (1943), Effect of electrical stimulation on atrophy of denervated muscle. *Arch. Neur. and Psych.*, 49, pp. 802–807.

von Schelling, H. (1944), Gedanken zum Weber-Fechnerschen Gesetz. *Abh. Press. Akad. Wiss. Math.-Nat. Kl.*, 5.

Wadley, F. M. (1946), Incomplete-block design adapted to paired tests of mosquito repellents. *Biometrics*, 2, pp. 30–31.

Wald, A. (1947), *Sequential Analysis.* John Wiley and Sons, New York.

Wiener, N. and Rosenbluth, A. (1946), The mathematical formulation of the problem of conduction of impulses in a network of connected excitable elements, specifically in cardiac muscle. *Arch. Inst. Cardiol. Mexico*, 16, pp. 205–265.

—— (1948), *Cybernetics.* John Wiley and Sons, New York.

Williams, C. B. (1947), The logarithmic series and its application to biological problems. *Jour. Ecol.*, 34, pp. 253–272.

Wilson, E. B. and Worcester, J. (1945), The law of mass action in epidemiology. *Proc. Nat. Acad. Sci.*, 31, pp. 109–116.

Wood, E. C. and Finney, D. J. (1946), The design and statistical analysis of microbiological assays. *Quart. Jour. Phar. and Pharmacology*, 19, pp. 112–127.

Yates, F. (1948), Systematic sampling. *Phil. Trans. Roy. Soc.* London. A, 241, pp. 345-377.

DISCUSSION FOLLOWING PAPER BY W. G. COCHRAN

J. W. Hopkins. I am sure that we are all extremely grateful to Professor Cochran for his most informative review of major areas of current research in biometrical theory.

As we all know, biometry is very much of an applied discipline. We thus have to recognize the fact that while the development of this discipline must of necessity be undertaken by persons of advanced mathematical attainments, the discipline itself is employed not only by these persons themselves but also in a variety of ways by a much more numerous and heterogeneous corps of applied workers. From the viewpoint of these latter individuals therefore, the developments that Cochran has outlined may be appraised in two ways: firstly, with reference to the extent to which the *corpus* of current biometrical theory covers the situations with which the applied worker has to deal; and secondly with reference to the extent to which the postulates or axioms from which the available theory is derived are in accordance with biological actualities.

As Cochran has mentioned, some unevenness of development is inevitable and, depending upon his particular requirements, the applied worker may find the extent of current theory more than adequate, reasonably adequate or inadequate. I should like to illustrate this with a few specific instances from my own experience, which is that of a biologist turned statistical practitioner in the laboratories of the Division of Applied Biology, National Research Council of Canada, where a somewhat diverse programme of laboratory experimentation in the fields of animal and plant physiology, bacteriology, fermentation processes, utilization of vegetable fats and oils and food technology is in progress.

The basic principles of sound statistical design, namely the elimination from experimental comparisons of major or systematic components of heterogeneity in the experimental material and randomization of the remaining variability, are of course the same here as elsewhere. It is to be remarked, however, that the object of a good deal of applied or technological experimentation is not to test for the presence or absence of possible effects, but to estimate the magnitude, under various specified conditions, of effects whose reality has previously been demonstrated. Such experiments, usually dealing with measured quantities, are commonly factorial in nature, and interactions, even of the highest orders, cannot be assumed negligible. They may indeed be of more interest to the experimenter than those which the agriculturist is accustomed to term "main effects". In these circumstances, devices such as fractional replication cannot be used and, if the experimental material is heterogeneous, recourse must be had to designs of the incomplete block or split plot type. These of course have been available for some years, so here we have an instance in which current theory is more than adequate to our particular requirements. On the other hand the more flexible and generally less time-consuming nature of most laboratory, as compared to field, experiments enhances the desirability of a sequential or (bearing in mind computing requirements) a two or three-stage progression to the number of replications required to yield confidence intervals of pre-specified width, as suggested by the speaker.

As an example of an extension of theory occurring more or less contemporaneously with the development of an applied requirement, I might mention a paper by Skellam discussing the nature of the discrete frequency distribution generated by drawing successive samples from a population in which the probability of occurrence of the attribute in question is not fixed but is itself a chance variable, fluctuating from sample to sample in a random manner specified by a Pearson Type I frequency distribution with range 0 to 1. This paper also outlined a workable method of approximating from observed data to the maximum likelihood estimate of the two parameters specifying the distribution. It appeared in the *Journal of the Royal Statistical Society* just as we ourselves were being forced to the conclusion that a series of sampling inspection results for large consignments of boxed fruit, which we had been asked to examine, could be adequately interpreted in terms of such a compound binomial frequency distribution, but not in terms of a Bernouilli binomial.

However, in our applied work we also encounter some problems that the theoreticians have not yet resolved. One of these was mentioned by Fieller yesterday. It arose in an experiment to study the effect on the yield from a fermentation process of the addition of small amounts of an organic supplement. It was found that the regression of average yield y on amount of supplement X was representable by a curve concave downwards having a point of inflection as well as a maximum, the corresponding regression equation thus being a polynomial of the third degree in X. The object of the experiment was to estimate the optimal amount of X (which was done by equating the first derivative, with respect to X, of the regression equation to zero) and also to deduce a confidence interval for this optimum. This latter constitutes a generalization of the familiar situation in bio-assay work in which a confidence interval is required for e.g. the estimated dosage producing 50% response calculated from the reciprocal of a linear regression coefficient. As Fieller mentioned yesterday, he has made this extension himself, but had not published it as yet. If he had, we should have been saved some trouble and anxiety in working it out for ourselves.

We encounter another more formidable unresolved difficulty whenever we want to deduce confidence intervals for components of variance estimated by differences. We want to do this quite frequently in comparing the variability of successive stages of a multi-stage production or sampling process in different locations or under different conditions, when measurements can be made only at the final stage. I personally regard the lack of an exact basis for statistical inference in such cases as the prime deficiency in current biometrical theory.

Finally, a word about the characteristics of actual biological measurements in relation to theoretical assumptions of normality, homoscedasticity, independence of residuals and so forth. We all know, or at least strongly suspect, that these assumptions are seldom fully realized in practice. We also know, however, partly from theoretical analyses and partly from the results of sampling experiments, that moderate departures from these assumptions do not seriously invalidate for example the tests of significance commonly applied to observed differences between means or between variances. What is lacking is any considerable or comprehensive body of recorded information respecting the statistical characteristics of different types of biological measurements. I suppose that of the 900 members of this society, probably 800 are, like myself, by profession consumers rather than producers of statistical techniques. Some at least of these 800 might make a very worthwhile contribution to the development of biometry by examining quantitatively and recording the characteristics of their data. This might not only indicate the extent of any to which

caution should be adopted in applying the customary statistical procedures; it might also provide mathematical statisticians with some specific points of departure for those unattractive studies of the effects of non-normality, non-linearity, etc., advocated by Cochran, if the need for these is demonstrated.

C. Gini. Si le rapport du prof. Cochran, avait pour tjtre *De quelques recherches récentes dans le domaine de la Biométrie*, il n'aurait aucune observation à faire. Mais, pour donner un tableau fidèle de l'état actuel de la Biométrie, il pense que le prof. Cochran aurait dû tenir compte de beaucoup d'autres recherces, dont Gini se borne à citer les quelques-unes qu'il se rappelle en ce moment.

1. Les recherches de Biotypologie. Ellesont donné lieu à des véritables écoles en Italie, en France et en Allemagne, trouvant expression en des revues scientifiques spécialisées, et ontcontribué au progrès de la clinique médicale, de l'endocrinologie et de la psychiatrie. En Italie, on a étudié aussi largement la mortalité et la fécondité différentielle des divers biotypes, qui a bien d'importance pour l'évolution des caractères des populations.

2. Les critériums pour juger de la normalité des proportions des hommes. C'est sur eux que se base le recrutement militaire. Le prof. Gini a pu démontrer que le critérium usuellement adopté porte à des résultats erronés dans le 20% des cas et a proposé des critériums plus rationnels qui avant la guerre avaient retenu l'attention du gouvernment belge.

3. Les recherches sur les jumeaux uniovulaires et biovulaires. Elles ont permis d'illustrer plus profondément, et sur des données précises, l'importance relative que l'hérédité et le milieu ont dans la détermination des caractères d'une génération.

4. Les études sur le rapport des sexes dans les naissances et les combinaisons des mâles et des femelles dans les familles humaines et dans les convades des espèces pluripares. Le prof. Gini lui-même a consacré un livre à ce sujet, il y a plus de 40 ans, et les recherches ont été reprises récemment en Italie et en France.

5. Les recherches sur les courbes du développement individuel. Bien des formules ont été proposées pour les décrire. Les recherches concernent non seulement l'organisme total, mais aussi les différents organes, et non seulement l'espèce humaine, mais aussi plusieurs espèces animales et végétales et ne se limitent pas à le période de la croissance, mais concernent aussi celui de la décadence sénile. Aussi sur ce sujet il y a des revues spécialisés.

6. Les recherches sur la taille des jeunes genes et en particulier des recrues militaires sur des données qui, dans quelques pays, s'étendent à des périodes très longues. En Norvège on a pu démontrer que l'augmentation de la taille a atteint, à l'âge militaire, pendant deux siècles, 12 cm. dont en calcule que la moitié soit due à une plus grande précocité. En d'autres pays, pour des périodes moins longues, le rythme de l'accroissement a été plus rapide, même 4 fois plus rapide. Les données dont on dispose concernent la plupart des populations européennes, celles de l'Amérique du Nord et quelques populations aussi de race jaune.

7. L'assimilation physique des immigrés. Les résultats de Boas, qui ont eu tant de retentissement parmi les hommes de science et n'ont pas manqué d'influencer la politique migratoire des Etats Unis, ont été confirmés par d'autres recherches faites en Europe sur les immigrés juifs de Berlin. Des investigations de plus large envergure ont été entreprises sur des différents groupes ethniques par le Comité

Italien pour l'étude des problèmes de la population. La dernière communication à ce sujet a été présentée en 1948 au Congrès International de Génétique de Stockholm.

Pour ce qui concerne les rapports numériques entre espèces en concurrence, à part les travaux de Volterra et de Lotka rappelés dans le rapport du prof. Cochran, bien d'autres ont été faits, avant et après, et leurs résultats ont été enfin encadrés dans une théorie compréhensive du développement des populations exposée en 1941 à Stockholm et publiée l'année suivante dans la Revue de l'Institut International de Statistique.

Les résultats sur la distribution des individus d'une espèce, dont le rapport parle aussi, font rappeler les études statistiques de M. Yule sur la distribution des espèces du même genre faites pour contrôler la théorie du Dr. Willis, tandis que ceux concernant les expériences cliniques font rappeler les recherches du Dr. Brownlee sur l'épidémiologie de l'espèce humaine que le prof. Greenwood et ses collaborateurs et successeurs ont continuées et étendues aux espèces animales dans leurs investigations d'épidémiologie expérimentale.

Les études de MM. Yule, Brownlee et Greenwood ont été faites en Angleterre; les recherches de Boas, ainsi qu'une partie très importante de celles sur le développement individuel, en Amérique par maints auteurs, entre lequels le prédécesseur du prof. Cochran, notre regretté collègue le prof. Pearl. Des autres recherches mentionnées, la plupart a été faite dans l'Europe continentale et une bonne partie en Italie. Si le rapport en avait tenu compte, la littérature qui le suit—dans laquelle pas un seul auteur Italien n'a été mentionné—aurait été beaucoup moins incomplète.

Le prof. Gini remarque que la coopération que la Société de Biométrie se propose, conformément aux voeux du Dr. Teissier et du prof. Cochran lui-même, exige de tirer profit des contributions de tous les pays.

Il pense aussi qu'il est bon de ne pas se borner aux résultats des dernières années, mais qu'il est nécessaire de tenir compte des études passées, car la Biométrie, loin d'être une science moderne, est la plus ancienne des branches de la statistique appliquée. Le fondateur de la statistique moderne, Adolphe Quetelet, s'était bien occupé des rapports entre climat et croissance des plantes et avait étudié la périodicité de la floraison ainsi que de la naissance et de la chute des feuilles, avec des résultats du plus grand intérêt qui mériteraient bien de n'être pas oublies. John Graunt, qui est regardé comme le fondateur de la statistique investigatrice, l'a été en réalité dans le domaine de la démographie; mais il n'était pas encore né quand Santoro de Capodistria donnait à l'Université de Padoue, pendant de longues années, des cours qui pouvaient bien être dits de Biométrie, dans lesquels il exposait les régularités que, moyennant d'ingénieux instruments, il avait mis en lumiére dans le domaine du poids et du métabolisme, ainsi que dans celui des pulsations du coeur, qui ont servi de base à l'école d'Iatromathématique qui cé s'est développée ensuite en Angleterre. Mais le développement de la Biométrie est bien plus ancien, car dès le siècle IX a.J. non seulement on avait systématiquement réalisé, dans l'Inde, des mensurations des différentes parties et des différents organes du corps humain, et étudié leurs rapports. en donnant lieu à une véritable Somatométrie, mais ces mensurations servaient aussi de base, outre qu'à l'art de la danse et à l'astrologie, aux représentations artistiques et à la médecine, devançant de la sorte de 27 siècles les théories de Quetelet sur l'homme moyen et celles des écoles modernes des biotypologistes.

J. B. S. Haldane. I wish to draw attention to three fields of population study not mentioned by Professor Cochran.

1. Paleontology. (a) Discrimination of species, especially of invertebrates. Two species may be quite indistinguishable on the basis of linear measurements or their ratios. But when account is taken of the fact that the regression of one measurement on another is non-linear, it is possible to determine characteristic parameters with great accuracy. (Kermack). (b) The absolute rate of evolution of a linear measurement or a ratio of two measures can now be measured. Typical rates of increase in horses and dinosaurs are about 4×10^{-8} per year (Pirlot and Haldane).

2. Animal demography. (a) Varley and others have obtained life tables with causes of death for insects, especially in galls (b) Blair, Hacker and others by repeated trapping of marked animals, have studied migration and consanguinity in small rodents. For example in a *Peromyscus* species about 10% of all matings were between brother and sister or parent and offspring.

3. Histology. Cells are being systematically counted in solid organs as well as blood. Among the results obtained is that of Boivin's school on the constancy of the thynucleic acid content of nuclei in a given species.

All these lines of work require their own statistical theory, with special references to their peculiar sources of error.

T. Kemp. As a biologist working in Human Genetics and a student of the late Professor Westergaard, I would emphasize that the sine qua non is the collecting of the primary observations of biological experiments in the right way. Less important is the way you treat your results statistically. New methods in statistics have been developed which increase considerably the assistance statistics is able to give to biology. Still the way the observations or experiments are done remains of primary importance and their statistical treatment will always be secondary. From this point of view, I feel, a useful cooperation of biology and statistics in biometry will be possible.

A. Rapoport. As a representative of the Chicago group of mathematical biologists, I am grateful to Professor Cochran for his excellent survey of the aims and methods of mathematical biology and for his plea for a broad concept of biometrics so as to include these aims. I am likewise indebted to Professor Rasch for emphasizing the fact that quantitative biology should go beyond the business of fitting empirical curves to equations; that it should actively seek a *rationale* which would explain the emergence of quantitative data, distributions, etc.

The Chicago group of mathematical biologists is oriented in just this way. They view their science as a deductive discipline, dedicated to the drawing of necessary conclusions from sets of assumptions. They are thus faced with a two-fold task: the choice of assumptions and the derivation of conclusions. The second part of the task is cut and dried. Its difficulties are entirely mathematical and its limitations are imposed by the frontiers of mathematical knowledge. Thus in developing their methods the mathematical biologists are often driven to the edges of explored territory in mathematics.

Much more intricate is the first part of the problem, that of choosing assumptions. The success of the mathematician in biology depends on the soundness of the biological theories presented to him by the biologist. But it is here that the mathematical biologist allows himself a somewhat sanguine attitude. He often bases a theory on a model which he knows to be quite inadequate. Sometimes no better model is available; at other times he is guided by mathematical expediency. He believes, nevertheless, that he is amply justified. The pronouncements of his science, as those of any deductive discipline, are never in the form "this is so" but always "if this were so, then this would have to be so." In other words, mathematical biology is committed to the development of theoretical frameworks in which assumptions based on future observations can be poured and from which their implications can be deduced.

The development of a rationale for biological events is the key task of mathematical biology. In this sense the development of a bad theory is better than no theory, because often the development of an inadequate theory provides insight into how it can be improved. To take a much cited example, it was necessary to treat gas molecules crudely and naively as billiard balls before more refined theories of gases could arise.

Finally I wish to express my hearty agreement with Professor Cochran's remarks concerning the hopeful indications of stochastic methods in mathematical biology. Since our first meeting two years ago, these methods have received increasing attention from nearly all of our workers, notably in Rashevsky's theories of the growth of cell populations and of rational thinking, in Opatowski's treatment of some problems of radio-biology, in Landahl's work in discrimination theory and in the probabilistic approach to the theory of neurol nets as recently developed by Landahl, Shimbel and myself.

In closing, let me again emphasize what I am sure is the feeling of the whole Chicago group that collaboration between mathematical biologists, biometricians and statisticians can be of extreme benefit to all concerned and that this Society is our natural meeting ground.

G. Darmois. Je veux seulement dire en quelques mots l'importance considérable des processus stochastiques en biologie, comme l'a indiqué, si justement le Prof. Cochran. Ma conviction est que toute la biologie est réelement le monde des fonctions aléatoires; les êtres vivants tels que nous, par exemple, se développent comme des fonctions aléatoires. Bien entendu, ce développement graduel d'une variable aléatoire peut être considéré comme se présentant déja dans les recherches de Laplace, et aussi dans celles de Kapteyn, dont vient de parler le Dr. Rasch. La théorie des fonctions aléatoires doit donc trouver dans la biologie un champ d'applications extrêmement fécond, et l'avantage sera que ces descriptions présenteront en même temps, comme celles de Kapteyn, une possibilité d'explication.

Other participants in the discussion included P. C. Mahalanobis and G. Rasch.

TEACHING AND EDUCATION IN BIOMETRY.

M. S. Bartlett

University of Manchester, England

1. *The definition and scope of biometry.*

It is with considerable diffidence that I have accepted an invitation to speak at this conference on teaching and education in biometry. My hesitation is primarily because biometry, however widely or narrowly we interpret it, is in the first place a branch of biology, and those of us who are professional statisticians must be careful to maintain a proper perspective. If we do not, we run a risk of attack comparable to that made on the early biometricians; for example, when Bateson in his Presidential address in 1904 to the British Association for the Advancement of Science,* speaking on genetical heredity and variation, asserted: "Operating among such phenomena the gross statistical method is a misleading instrument; and applied to these intricate discriminations, the imposing Correlation Table into which the biometrical Procrustes fits his arrays of unanalyzed data is still no substitute for the common sieve of a trained judgment". . . . However, Bateson's strictures were coloured by a controversy since resolved; and the advances in biometry due to those like Karl Pearson and R. A. Fisher, who, whatever their many other qualifications, may certainly be called statisticians, give us some right to speak. I am moreover indebted to my colleague at the University of Manchester, Professor Eric Ashby, for indicating to me in conversation his own viewpoint as a biologist, and I hope in consequence that my discussion will not be unduly biased.

I am not sure what is the "official" definition of Biometry or Biometrics accepted by our Society but I assume from the reference to "quantitative biology" in an early prospectus that the Society takes a wide view of its scope and does not necessarily identify this with the particular investigations of biological variation, co-variation and quantitative inheritance by Karl Pearson and other early biometricians, however important a nucleus these investigations constituted. Clearly we must include the further researches with which R. A. Fisher's name is so

*See (1), p. 578 (I owe this reference to a popular article on genetics and biometrics by Julian Huxley (2)).

closely associated: the development of statistical theory, methods of experimentation and analysis in agriculture, genetics and in general biology; we must include bio-assay and the study of the growth and competition of populations,—which brings us to quantitative ecology (including human ecology),—and lastly we must not forget the study of the growth and form of individual organisms, (as described, for example, in (3)).

There is to me a logical similarity perhaps worth noting in the relation of biometry to biology on the one hand and of econometrics to economics on the other. Professor Gerhard Tintner, when he was discussing the scope of econometrics recently, pointed out that this subject must not be confused with economic statistics or mathematical economics, both of which are used in econometrics in an attempt to synthesize observation and theory and build up in economics realistic and quantitative theories supported by observation. If we substitute "biology" for "economics", and so on, we have the logical relations of biometry to biology, biological statistics, and mathematical biology, (and if C. I. Bliss intended any identification in his phrase "biometry, or statistics applied to biology" (4) I would query his definition). Of course in contrast with economics there is the great advantage in biology both of the experimental method and of observational material of much greater abundance.

After this amplification we may return to our starting point and think of biometry briefly as quantitative biology. One incidental point made to me by Professor Ashby is that it is convenient to exclude from our definition purely biological experimental techniques,—for example, techniques in preparing slides for the microscope to study cell and tissue growth,—not because they are not concerned with biometry and measurement, but because they are better classified and included by the biologist as part of his general laboratory training, in contrast with the rest of biometry which relies, at any rate in part, on mathematical and statistical concepts.

2. *Biometry and the biological student.*

Coming now to the teaching of biometry, we shall find it convenient to think in terms of various levels; the mathematical statistician will certainly be involved with the advanced stages; but it is further convenient and, in view of my above remarks, logically appropriate to begin with the more elementary levels of concern to all biological students. Before thinking specifically in terms of biometry, it may be useful to quote some conclusions by K. Chapman in a recent article (5) on the place of biology generally in university education:

"(1) Biology is entitled to an important place in university education because of its unique and present-day value in including an 'evolutionary attitude' to natural and human phenomena. (2) It is of value also, not only as an important 'practical' subject through its connections with hygiene and human physiology, but also as a subject basic and fundamental to sociological and psychological studies as well as medical studies on human beings. (3) The content of university course in biology, more especially in zoology, needs reconsideration in some universities here* and in Commonwealth countries abroad; both from the practical and the academic, philosophical point of view, it is worth considering whether the quality and quantity of ecological, taxonomic and genetical studies might be increased. This is particularly so with regard to education in appreciating evolutionary phenomena and problems. (4) At the same time, the factual content of university courses in biology should not be allowed to become great enough to obscure its essential function as illustration of general principles. (5) Guidance in appreciation of the 'synthetic' as well as the 'analytical' content of scientific work and thinking should be given in this as in other sciences".

It is not my concern nor my place to discuss these proposals in general, but merely to note the references to more "ecological, taxonomic and genetical studies", to "an evolutionary attitude" and "appreciation of the synthetic . . . content", and to make the obvious remark that in such studies biometry, biological statistics, and the quantitative approach to biology will play an important role.

The courses taken in the botany department at Manchester include what seems to be an admirable attempt in this direction, biometry, in the wide sense in which I am using this word, being covered at the undergraduate level by the following topics:

(a) general introductory (including laboratory) work on growth rates, including simple quantitative ideas such as logarithmic increase; (b) genetics, including elementary quantitative genetics; (c) a course in quantitative ecology, including field work during the summer in, for example, measuring plant densities, and constructing ecological maps; (d) simple statistical concepts, which arise naturally in connection with the content of such courses, e.g. the treatment of frequency distributions, dispersion, skewness, bimodality, etc.

While there is not any systematic statistics course, it should be remembered that these courses are for all botany students, not just for specialist biometricians, and it would seem exceedingly reasonable to allow the understanding of statistical and quantitative concepts by the student to grow gradually with the need of proper analysis of his own biological material. But, as Professor Ashby points out to me, and as is indicated also by Chapman in his article, such courses are by no means

*That is, in Great Britain.

routine in all universities, and (c) in particular is probably rare among undergraduate courses.

There has been some discussion in recent years as to whether statistics should be taught in schools (see, for example, (6)), and some draft mathematics school syllabuses have included elementary statistics topics. But apart from the obvious need to have university syllabuses revised first, doubt has been felt by mathematicians as well as by others concerning how far somewhat abstract ideas can be "got across" at the school level. The possible answer to this difficulty is evidently to see that the statistical ideas are absorbed by means of actual sampling and study of statistical populations (e.g. of a biological kind), a hobby which I have been told can be made as fascinating for schoolboys as their indefatigable collecting of bus-tickets or of the names of locomotives.

However that may be, let us return to the university and to the biological student as undergraduate. We have to consider what more should be done for the student who wishes to specialize further in biometry and biological statistics. Again I can speak with most knowledge of the existing arrangement in my own university (and I hope there will be an opportunity to hear at this discussion how far arrangements differ elsewhere), where a number of statistics courses at various levels are given in the Science Faculty, somewhat on the lines recommended by the Royal Statistical Society (7). These begin with a very elementary and non-mathematical course of about 12 lectures, of most interest to ordinary or general degree students in science. The next stage, a "subsidiary course" consisting of an introduction to statistical theory, statistical methods and a practical class, is largely attended by honours mathematics undergraduates as part of the subsidiary subject they require to study in addition to their main degree work, and as an essential preliminary to taking statistics as an option in their main degree work itself. However, it is also intended to be available for any students with an elementary knowledge of mathematics (algebra, and elementary differential and integral calculus) who wish to take the course, whether they are physicists, economists, biologists, etc. Such dual or multi-purpose courses are not altogether ideal, and I frankly regard this one as a provisional arrangement which may have to be modified in the light of experience. But in view of its intended relevance for the biological student who has some mathematical background and wishes to begin to study statistics seriously, a brief outline of its syllabus is added:

(a) *Introduction to theory.* Historical survey; elements of probability theory; observational and theoretical frequency distributions; binomial and Poisson

distributions; averages and moment-generating functions, elementary sampling theory; properties of normal or Gaussian distribution; simple correlation, regression and least-squares. (b) *Statistical methods.* Description (without proof) of various standard statistical techniques, such as multiple regression and least-squares, including numerical procedures, orthogonal polynomial fitting and Fourier analysis; analysis of variance and principles of experimental design; correlation analysis; principles of sampling. (c) *Practical* ss. Examples illustrating topics dealt with in lectures.

It is suggested, in the recommendations by the Royal Statistical Society already referred to (7), that a course (presumably of about the above standard) should be a normal and not an exceptional course for biological and other students: "We believe that it is desirable that this second type of course should be taken by most students of the social and biological sciences in all universities, as part of the normal curriculum for a first degree. We realize that this may be regarded as a revolutionary proposal, and that there are great difficulties in adding further work to existing programmes, but it is impossible to escape the conclusion that the scientific worker in the social and biological fields who has not a sound grasp of the statistical approach and its simpler methods of application, is seriously handicapped when appreciating quantitative laws and attempting the interpretation of observational data."

One familiar difficulty (cf. 8) is the rudimentary or forgotten mathematical knowledge of biological students (and of course of others such as economics students), and Professor Ashby has made to me the rather interesting suggestion that instead of a statistics course with perhaps a separate revision course in mathematics for those who require it, a more "portmanteau" though elementary course should be designed for biologists, in which statistics and mathematics should be combined. He also thought that it would perhaps best be given by a biologist. The last point touches the general controversy on how far mathematical statisticians should undertake the statistics teaching for students in different departments, a question which I personally think has no "best" answer and is a matter for compromise (cf. (7), and my discussion in (9)). Both on this general problem and on the particular suggestion of a combined mathematics and statistics elementary course, the opinion of this Conference would be welcomed. I have therefore ventured to list a number of mathematical and statistical topics which seem relevant, though not all of them could be accommodated in an elementary course. The complete list goes even further than the "subsidiary course" syllabus above, and, intended in particular for the biometrician, would certainly provide him with a very useful technical background.

Draft syllabus of combined mathematics and statistics course for biologists.

(Note: For convenience mathematics is listed first, but the entire order is of course a matter of expediency. Whether starred items are included depends on the length and comprehensiveness aimed at, and certainly items doubly starred would not be included in an elementary course. Examples would be drawn from biological sources; e.g., differential equations from problems of growth, frequency distributions from ecology, analysis of variance from agriculture, maximum likelihood from genetics, probit analysis from bio-assay. The course would include practical work).

(a) *mathematical technique*

binomial theorem (for positive, *fractional and *negative indices); *limits and convergence; geometric, *exponential and *logarithmic series; *solution of simultaneous linear equations (numerical method); functions of one variable, differential coefficients, definite integrals; interpolation, finite differences, *numerical integration, numerical solution of equations, *Taylor's series (symbolic derivation), *elementary difference and differential equations; **functions of two or more variables, **partial differentiation, **definite double integrals; Cartesian geometry of two dimensions,—line, conics referred to simplest coordinate axes; *Cartesian geometry of three dimensions, —plane.

(b) *statistical technique*

observed frequency distributions in one variable, graphical representation, parameters of location, dispersion and skewness; theoretical distributions, binomial, Poisson, *negative binomial and **logarithmic, normal; simple reduction and analysis of data, averages and dispersion in small samples; sampling variation, standard errors and tests of significance (contingency tables and chi-square, t test, analysis of variance, elementary principles of experimental design); simple correlation and regression, **normal law in two variables; *multiple regression, *method of least squares, *analysis of covariance; **method of maximum likelihood; **probit analysis and other transformation methods; **planning sampling surveys.

3. *Biometry and the mathematical statistician.*

It is being increasingly recognized that the ever-increasing front to the wave of scientific discovery is making suitable co-operation essential; otherwise, further advance may well disintegrate into incoherent independent zones of sterile activity. As a *Nature* editorial (10) on "Training the Research Worker" remarked: "In spite of all that has been accomplished by individual genius, the sound progress of a body of science requires even more the individual who combines with outstanding intellectual urge of a personal nature the ability to co-operate with colleagues of equal attainments and conditions; the obdurate lone worker can, in fact, impede the evolution by a group of scientific men of a joint integrated programme of research".

Such co-operation is especially important in the development of statistical applications, because so few of us are able to combine both a deep practical knowledge of any field with a comprehensive theoretical

knowledge of modern statistical technique. This of course raises the controversial problems already hinted at, whether in teaching or in research; but for my immediate purpose it is sufficient to note the useful contributions to biometry which mathematical statisticians can make to justify a brief survey of their biometrical education also.

Before this can be done, we have to decide what broad quantitative and statistical aspects of biology are important at the present time, for only then can we judge on the efficiency of current teaching practice. It is quite permissible in this task to classify statistical concepts and techniques from their abstract content rather than in connection with their particular biological field of application, for it is especially the function of the theoretician and statistician to examine the general structure of the wood rather than the individual trees, however beautiful and foliaceous these may be.

Historically, we may remember that statistical investigations began with the empirical study of human populations, though soon the independently developing probability calculus was brought in to clarify theoretical statistical concepts. Great advances in the present century in the theory of statistical inference have been an important factor in getting recognition of mathematical statistics as a respectable subject of university study. This, combined with the feasibility of teaching statistical methodology as a particular but important and well-defined branch of scientific method, has led most modern "mathematical statistics" courses to concentrate on this aspect. They may, as a glance at prospectuses of American and British universities will show, include a dazzling array of important branches of the subject of statistical inference,—for example, statistical estimation, testing statistical hypotheses, experimental design, multivariate analysis, sequential analysis,—touching on their abstract side the fundamental problems of scientific induction and on their practical side the most varied fields for their application. It is certainly not my desire to decry the immense value of these courses, but I would like to stress another complementary aspect of statistics of equal if not greater importance, the theory of statistical populations or aggregates, and of "statistical laws". It is this aspect which led Erwin Schrödinger (11; cf. my remarks in 9) to refer to the "Statistical Law in Nature"; it is imperative for physicists and biologists alike to grasp the nature of statistical phenomena.

Now this aspect can even less easily be taught divorced from its particular field, for it is in fact the vital place which the statistical laws occupy in such a field, physical or biological, that give them their importance; and this means a fairly detailed technical study of the field

itself. Thus it is not usual for statistical mechanics to be taught as a branch of statistics; while, since biologists are not on the whole as mathematically-minded as their physicist colleagues, it is unusual for the statistical aspects of evolution, as developed by R. A. Fisher, J. B. S. Haldane, Sewall Wright and others, to be taught at all.

Nevertheless, I believe that in the preliminary education of the mathematical statistician, in particular if he is to be interested in biometry, this aspect should if possible be represented. It may help to provide a link with macroscopic vital and other social statistics, more familiar in the applied statistics courses of other departments, but interest in which should not be confined to the medical statistician, economist and actuary. Various topics beginning to be included in theoretical statistics courses are, I think, indicative of this development and synthesis of statistical theory. Thus we find more attention being paid to composite populations, whether in the study of polls and sampling surveys, in the theory of discrimination and classification of individuals from different groups or in the theory of frequency distributions of numbers of individuals in different species (of animals or of words; see, for example, 12). We find further study of how populations grow and distributions develop, so that on the one hand the classical actuarial theory of population growth developed by Lotka and others becomes extended to include the study of stochastic effects; and on the other, particular frequency distributions like the negative binomial and logarithmic are found to emerge from more than one causal mechanism. These last topics are examples of "stochastic processes", the further theoretical study of which will broaden still more the understanding of how particular statistical laws and distributions arise,—for example, under what conditions the numbers of interacting species oscillate, the chance of a particular type of population becoming extinct, or the possibility of a recombination fraction in genetics of over 50%. It is significant that in a recent symposium (12) on stochastic processes, while a paper by D. G. Kendall on population growth and to a lesser extent my own contribution, would be of most interest to biometricians, the same mathematical technique was used by J. E. Moyal in his study of stochastic processes in physics; this completely justifies the general study by the mathematical statistician of these processes, as well as their detailed study in relation to their particular fields of origin.

To return at last to the mathematical statistician as student, after this realisation of all that he ought to know,—as a mathematical undergraduate he will still have done well to have learnt the main principles and techniques of statistical inference, with perhaps an introduction to

the theory of stochastic processes, but with what is more important as a balance to his rather abstract studies some training from practical classes in statistical method. He will then still be far from being a biometrician, but is perhaps in a reasonable preparatory state, though from very different stock from the biological student.

4. *Biometry and the postgraduate student.*

With either approach there seems no doubt that the main education of the genuine biometrician has still to begin, and must inevitably be a postgraduate one, as before then his other degree work is too time-absorbing. To some extent an attempt has been made in English universities, for example, at Cambridge, Oxford and Manchester, to ease the difficulty of the transition by providing the opportunity, if desired, of taking a post-graduate Diploma in Statistics. This falls into line with the increasing recognition of the amount a scientific worker now has to learn before he is ready to begin research, a point perhaps more neglected by English universities than their American counterparts, where regular post-graduate courses are a recognized part of higher degree work. This neglect has been in particular noticeable in statistics, owing to the general neglect of this subject at the undergraduate level,* and the Royal Statistical Society has also recently initiated a Diploma examination for its Fellows (14).

At Cambridge the Diploma is at present usually taken in Mathematical statistics, with a selected field of application in which the candidate must do practical statistical work and which (in the present context) could be Agriculture or Biology, though the separate Cambridge Diploma in Agricultural Science has the option Statistics available as an alternative for the would-be agriculturist; there is not yet, I believe, any corresponding alternative for other biologists. At Oxford the candidate must take in addition to general statistical theory and practice two among a number of special papers, of which, while he may substitute for one of them if he prefers a paper on the Mathematical Theory of Statistics, the following two are the most relevant, (the Diploma syllabus has been published (15) and the two appropriate sections of it are appended for reference):

"*Design and Analysis of Scientific Experiment.*

Analysis of variance and covariance.
Estimation and inference from experimental results.
Elementary biological assay.

*Perhaps the most noteworthy exception is at University College (London University), where for many years it has been possible to take a B.Sc degree in statistics.

Transformations for non-normal data.
Principles of experimental design; randomization, replication, orthogonality, balance.
Knowledge of standard designs and methods of analysis of results.
Factorial experimentation; confounding.
Elementary problems in statistical analysis of genetical data.
Application of these topics in special fields of scientific research.

Sampling Methods in Scientific Research.

Sample estimation for discrete and continuous variates.
Precision and bias.
Random and systematic sampling.
Stratification.
Double sampling.
Planning, execution and analysis of sampling surveys in agriculture, forestry, biology and psychology, (choice of sampling unit and strata, intensity of sampling, etc.)."

At Manchester the Statistics Diploma can be taken either by the mathematical statistician, who must in addition to further general study in statistics choose, as at Cambridge, a practical field of application, which for the potential biometrician could be "biology and agriculture", or by (again in the present context) the biologist, who in addition to further practical and theoretical study of statistics in relation to his own subject, would supplement further his general statistical education.

I have included these references to the English postgraduate Statistics Diploma in the hope that they may also be of interest to those present from other countries; for though some of the reasons for its institution are no doubt peculiar to my own country, the general problem it tries to solve of introducing a student to a new field of study outside his traditional paths, (in particular, a mathematician to biometry or a biologist to mathematical statistics), is universal.

The need for the dual approach from mathematics *and* biology has been also stressed by Professor Cochran in the article (8) already referred to on the training of statisticians for biological work. He says in his summary: "For the would-be professional statisticians the training should include mathematics, mathematical and applied statistics and some broad, fundamental courses from the applied field which the student selects. The training should be sufficiently flexible to cope with aspirants from both the mathematical and biological fields and should culminate in the conduct of research".

It will be remembered that I have myself distinguished biometry from biological statistics, but the training of biometricians and biological statisticians will of course have much in common. However, even after the postgraduate student's formal training, it is evident that only a very

partial metamorphosis will have been reached. The best training and education in biometry, as in other subjects and as Professor Cochran mentions, must ultimately be, when the preliminary training has been complete, by research experience,—either in one of the biology departments at a university or at a biological research station, a famous example of which in England is the Rothamsted Experimental Station where R. A. Fisher carried out so much of his biometrical and statistical research. In such research, moreover, it will be surprising if some co-operative effort is not required. The habit of co-operation and learning from one another can also, I think, be much promoted by appropriate seminars, meetings and conferences, of which the present international conference is one. Thus while the final emergence of a fully developed member of the species is never likely to be a common event, let us at least hope that this co-operative habit will enable mere statistically-minded biologists and biologically-minded statisticians to mimic successfully that much rarer creature, the perfect biometrician.

REFERENCES

(1) *Report of the British Association for the Advancement of Science; Cambridge meeting*, 1904.
(2) Huxley, Julian. Half a century of genetics. *Sunday Times*, July 10, 1949.
(3) *Essays on Growth and Form*, edited by Le Gros Clark, W. E., and Medawar, P. B., Oxford, 1945.
(4) Bliss, C. I. Biometry in the service of biological assay. *Ind. and Eng. Chem.* 13, 84, 1941.
(5) Chapman, K. The place of biology in university education. *The Universities Review* 21, 176, 1949.
(6) The teaching of statistics (a discussion opened by Dr. J. Wishart). *J. Roy. Statist. Soc.* (*Series A*) 111, 212, 1948.
(7) Report on the Teaching of Statistics in Universities and University Colleges, *J. Roy. Statist. Soc.* 110, 51, 1947.
(8) Cochran, W. G. Training at the professional level for statistical work in agriculture and biology. *J. Amer. Statist. Assoc.* 40, 160, 1945.
(9) Bartlett, M. S. Mathematical statistics and the universities. *Universities Quarterly* 3, 665, 1949.
(10) Training the research worker. *Nature* 163, 969, 1949.
(11) Schrödinger, E. The statistical law in nature. *Nature* 153, 704, 1944.
(12) Simpson, E. H. The measurement of diversity. *Nature* 163, 688, 1949.
(13) Symposium on Stochastic Processes (papers read by Bartlett, M. S., Kendall, D. G., and Moyal, J. E., on June 9th, 1949 and to be published with discussion in *J. Roy. Statist. Soc.* (*Series B*)).
(14) General Regulations for the 1948 and 1949 Certificate Examinations for Fellows and for the 1949 Diploma Examination for Fellows. *J. Roy. Statist. Soc.* (*Series A*) 111, 71, 1948.
(15) Diploma and Certificate in Statistics (Regulations), *Oxford University Gazette*, p. 628, April 21, 1948.

DISCUSSION FOLLOWING TEACHING AND EDUCATION IN BIOMETRY

M. S. Bartlett

W. G. Cochran. Since Dr. Bartlett has asked us to contribute different types of experience to the discussion, I should like to mention two courses given at Johns Hopkins University. The first is a course for medical students in the first year of their medical curriculum. The majority of the students expect to become medical specialists: only a few will ultimately devote themselves to research. Consequently the problem is to present those aspects of biometry that are most relevant as part of a general education in medicine. The available time, 36 hours, is divided into 12 hours of lectures and 24 hours in the laboratory.

The course is planned with the object that at its end the student should be able to read the medical literature more intelligently and critically, to use a few elementary techniques, and to realize when he needs further advice from a biometrician. The following topics are included: issues involved in collecting and appraising evidence; reasoning from rates and ratios; prognosis figures derived from long-time follow-up studies; the sampling variation of rates; the analysis of measurements; bioassay; the planning of medical investigations; and dynamic theories in medicine.

Considerable importance is attached to the laboratory work, in which the student handles a collection of problems taken from the medical literature. Stress is laid on the scientific issues involved rather than merely on details of computation. A varied selection of optional problems is included. Laboratory assistants (about 1 assistant to each 10 students) are present to discuss the problems with the students.

As the value of instruction in biometry is more widely appreciated, I believe that there will be an increased demand for courses of this type, in which the students are potentially neither biometricians nor research scientists in some branch of biology. To be successful, such a course must be constructed with the interests of the students in mind. How far it should differ from the type of course that one would recommend for the training of biometricians is a question that is not too easy to answer.

My second example is a course entitled "Statistical methods in epidemiology," and given jointly by the departments of epidemiology and biostatistics. The topics include an introduction to epidemic theory and an account of the methods used in the study of both acute and chronic diseases. In general, the role of the epidemiologist is to set forth the known biological facts and speculations about the disease and the ways in which it can spread, and to present the data that are to be the subject of discussion. The biometrician outlines the mathematical and statistical tools that can be used in analysis of the data. Both together discuss the validity of any assumptions made in the analysis, the biological meaning of the results, and the utility of this approach, sometimes in the form of a good-natured though heated debate. As might be expected, students have been very enthusiastic about the course. I mention this example because it is one that might profitably be tried with other branches of biology. It ensures a close integration of biometry with biology that is one of the aims of our society.

D. J. Finney. A problem that has concerned me for some years and one that has been mentioned several times this week is the use of simplified and approximate methods in statistical analysis. I suggest that the statistical teacher and consultant

need to give more thought to the part which these should play in statistical methodology.

I have in mind, for example, the analysis of rank orders instead of quantitative measurements, the use of the range for assessing variability instead of sums of squares, and graphical or other simplified methods for the analysis of biological assays. The aim of these is to save the user much of the arithmetic labour required in the methods generally recommended by statisticians. Unfortunately, however, there are usually theoretical objections or limitations to their application. In well-planned experiments, such a simplified analysis will often lead to essentially the same conclusions as would a more orthodox but more laborious method, but occasionally there may be serious differences. Even the experienced statistician may have difficulty in recognizing, on inspection of the data, whether or not a simplified method may safely be used, and he will naturally be very reluctant to teach his students or to advise his non-mathematical colleagues how to use one. To ignore the simplified methods is unsatisfactory and uneconomic, to advise their use is dangerous without careful qualifications, and at present we are seldom able to state the right qualifications. I believe that those of us who are teachers of biometry need to investigate all methods of analysis with a view to being able to state precisely their limitations and to assimilate the best of them into the main body of biometric methodology.

C. I. Bliss. In defining biometry, it is useful to distinguish between the formation of hypotheses, which is sometimes called biophysics or biomathematics, and the testing of hypotheses, which is in the domain of biostatistics and includes the design of experiments, problems of estimation and tests of significance. They are both biometry but they require somewhat different types of training on the part of both the instructor and the student. Courses tend to center about one aspect or the other, depending upon the instructor's capabilities and interests. Professor Bartlett has quoted my thumbnail definition of biometry in 1941, which reflects my own bias. My 1949 definition would be more inclusive and based upon that adopted by the Biometric Society in the first article of our constitution, even though the content of the courses I teach is still conditioned by my initial training and experience as a biologist.

F. Yates. I speak as a worker in a large Research Organization concerned with supervising working research statisticians, not with teaching. In training mathematically minded biologists or biologically minded mathematicians so that they will become useful working statisticians we should in the first place provide them with a battery of standard methods to meet recurring situations. It is very important that they should understand the underlying principles and recognize to what situations the method applies without qualification, thus, to take the analysis of variance, this can be applied to the results of randomised and replicated experiments: here a significant result gives evidence of a causal effect arising from the treatments. It can be applied to measurements on groups of animals receiving different treatments, each group being kept in a single pen: here significance may indicate pen and not treatment effects, e.g. disease in one pen. Again it can be applied to observational data, when causal effects cannot be assumed. The student must also be taught to recognize clearly the complications which prevent a standard procedure being applied without modification—danger signals must be erected—for example, non orthogonality in the analysis of variance. Finally the student, if of sufficient promise, may be trained in the art of developing new methods and modifying existing methods

to meet new and infrequently occurring situations. When he reaches this stage there is less difficulty for the teacher, as the student should be capable of thinking for and in large measure teaching himself.

The mathematical machinery is not necessary for the working statistician. For example, it is very important for him to know that the χ^2 distribution for n degrees of freedom is the distribution of the squares of n normally and independently distributed variates each with unit standard deviation. But one need not know how to derive the distribution mathematically, or integrate it.

The teacher must be prepared to assist his students in cutting away dead wood; theory and methods which have been superseded and are no longer of value.

We live in an instruction-book age. Statistics is a technique. There is great need for more specialized technical books. We must also teach students to use these books rather than attempt to provide them with new books in courses of lectures.

S. N. Roy. In the Indian Universities and the Indian Statistical Institute there is no specific course designed for the biometrician as such but a generalised statistical course which would help the biometrician up to a certain point. It could serve as the basis for additional courses for the biometrician as for workers in other fields. The general statistical course in the Indian Statistical Institute is unique, however, in that there is greater coordination between theoretical teaching, statistical computation and actual field work than in most other centers.

A second point concerns the importance of statistical inference in statistical teaching, whether for the biometrician or for the worker in other fields. Both in the limited technical sense and in the wider sense of the broad objectives of all statistical investigations, the theories of statistical inference might well permeate the whole of statistical teaching—and not be limited just to certain sections of statistical courses. This would help prevent the misuse of statistical methods by the average statistician although no amount of statistical teaching can be a substitute for common sense, intuition and aptitude.

A. Vessereau. Il faut que, dans l'enseignement, soit prise en considération la nécessité de former des équipes de biologistes et de statisticiens. Les étudiants en sciences naturelles ont souvent les plus grandes difficultés à comprendre et même à accepter le mode de pensée et de raisonnement qui s'exprime en calcul des probabilités et en statistique mathématique; dès le début de l'enseignement de la statistique, il faut donc beaucoup insister sur cette nouvelle position de l'esprit. Dans l'exposé d'une question de statistique mathématique, il faut tout d'abord préciser très clairement les hypothèses—indiquer même très rapidement la ligne suivie pour atteindre la solution—formuler cette solution en supposant à nouveau des hypothèses pour bien délimiter le domaine de son application—enfin passer aux méthodes pratiques d'application, avec de nombreux exemples sur lesquels les étudiants eux-mêmes auront à travailler. En aucun cas il ne faut se borner à donner des formules ou des recettes sans en préciser le sens.

Pour amener des mathématiciens à la biologie, il serait utile que, dans l'enseignement supérieur de base de mathématiques, une place soit donnée à des leçons destinées à découvrir les vastes horizons de la biologie à des étudiants qui n'en ont en général aucune idée.

Il est précisé que les considérations ci-dessus s'appliquent à l'enseignement français.

Other participants in the discussion included G. Darmois, G. M. Cox, M. H. Quenouille, F. Bernstein, E. van der Laan and L. Martin.

QUERIES

75 **QUERY:** We are attempting to biologically standardize thyroid preparations by means of an indirect measurement of increased oxygen consumption in mice. We employ a 6-point assay with three equally spaced doses of the standard and test preparations. The animals are injected daily for ten days. On the eleventh day they are placed in individual air-tight Mason jars and their survival time is determined. Potency is then calculated by factorial analysis.

I am enclosing the raw scores of a recent assay and also the analysis of variance table. When the variation between doses is reduced to single degrees of freedom, I find, in addition to a highly significant F value for

Survival Time in Minutes

Lot	Standard Preparation (Mg. Thyroid per Day)			Test Preparation (Mg. Thyroid per Day)		
	0.6656	1.664	4.16	0.6656	1.664	4.16
1	107	74	61	69	71	67
2	77	99	43	100	64	58
3	68	63	53	64	43	58
4	64	75	73	76	69	50
5	72	57	78	68	51	61
6	116	53	64	84	60	53
7	91	75	64	86	68	64
8	114	72	54	74	70	59
9	79	53	75	65	62	54
10	87	66	57	65	74	63
11	77	62	56	71	48	58
12	151	55	67	71	98	65
13	96	56	59	75	63	60
14	57	71	58	69	62	59
15	72	77	74	83	72	57
16	73	74	89	103	55	66
17	83	76	54	90	104	66
18	86	74	63	82	88	56
19	85	73	54	48	98	59
20	97	68	84	63	79	59
21	95	55	59	71	82	73
22	63	56	39	85	74	51
23	88	78	59	85	74	59
24	91	73	60	79	72	58

regression and a non-significant value for departure from parallelism, a significant (P = 0.05) F ratio for the "opposing curvature" term. I am not able to interpret this and have not been able to find anything similar to it in the literature.

I would especially like to know if this significant departure from linearity has any effect upon the final determination of potency and if it might indicate the accidental introduction of "bias" in the assay procedure.

Analysis of variance

Source of Variation	Degrees of Freedom	Mean Square	F
Preparations	1	473	2.59
Regression	1	10,106	55.40
Parallelism	1	412	2.25
Quadratic	1	104	0.57
Opposing Curvature	1	732	4.01
Error	138	182.4	

ANSWER: When a test of significance is used as a routine on a series of similar experiments, it is important to remember that, of effects really due to chance, one in twenty will give test criteria beyond the 5% significance level. If your assay is one of a series, and others have given no sign of significant "opposed curvatures", I should be inclined to dismiss the present occurrence as due to chance and to proceed with the estimation of relative potency in the ordinary manner. Belief in this explanation is encouraged by the fact that neither the parallelism nor the direct quadratic component is large enough itself to arouse any suspicion of invalidity. If the assay is an isolated experiment, I should still incline to this explanation, though I should state my conclusions with caution and should ask for a repetition of the assay if the result were very important.

If, on the other hand, this assay is one of a series in which some indication of invalidity has frequently occurred, perhaps sometimes in the parallelism or the quadratic component, I should be inclined to suspect the validity of the basic assay assumptions or of the method of statistical analysis. In this connection, you may note that your analysis implicitly assumes the variance of the responses to be the same in each of the six dose groups. When time of survival is used as a response, the less effective doses may show occasional instances of very large responses. Examination of your data, indeed, shows the variance to be considerably

larger for the two low doses than for the two high. Use of the reciprocal of the time of survival as a response metameter often corrects such heteroscedasticity, and I have therefore tried the effect of this on your data. In order to keep convenient whole numbers, I used the response metameter

$$y^* = 10{,}000/y$$

for each of the 144 responses, and obtained the following analysis of variance of y^*:

Source of Variation	d.f.	S.S.	M.S.	F
Preparations	1	865	865	1.3
Regression	1	37,683	37,683	54.8
Parallelism	1	852	852	1.2
Quadratic	1	84	84	0.1
Difference of quadratics	1	1,043	1,043	1.5
Between doses	5	40,527		
Error	138	94,832	687.2	
Total	143	135,359		

Not only are the differences in the variances within the six groups much less serious for y^* than for y, (as may be verified by subdividing the error sum of squares into six sets of 23 d.f.), but none of the variance ratios for parallelism, quadratic, and difference of quadratics is now large. This suggests that the reciprocal of time may be a more suitable measure of response than time itself in your assays of thyroid preparations. I certainly would not regard the evidence of one assay as sufficient to justify a change of metameter, but no doubt you have other sets of data that you can examine. In a paper published in the December 1949 *Biometrics*, I discuss the effect of choice of metameter on the results of an assay, and, in another paper in the same issue, N. K. Jerne and E. C. Wood consider the theoretical requirements for a good metameter. You may note that, for your data, the direct analysis of y gives a relative potency of 1.38 for the test preparation, with 95% fiducial limits at 0.93, 2.16. The analysis of y^* gives an estimate of 1.25, with limits at 0.84, 1.94. Thus from the point of view of any action to be based upon the assay, the two analyses give effectively the same results; I suspect that the second is a better representation of the real state of affairs than the

first, but I should want to see data from many assays before expressing a final opinion.

D. J. FINNEY

76 **QUERY:** Another point which came up in connection with frozen peas (Query 74, December, 1949) bothers me. As explained there, judges were given three dishes of peas, two from one temperature and one from another. We randomized the choice of the temperature to be duplicated. On separate examination we find that the judges score more correctly during the early weeks when the lower temperature was duplicated. This has also shown up in other experiments. Can you offer any suggestion as to its cause?

We cooked three 300 gram samples of peas of each lot at each tasting period, two from one temperature and one from another. Following cooking, the two samples from one temperature were mixed before taking samples for the judges. Could this procedure have influenced results?

ANSWER: Let me take your last question first, since the answer is "yes". Suppose that the average quality and the variability from sample to sample, including the effect of cooking, were the same for peas stored under condition *A* and condition *B*. After mixing there should be almost no difference between the two samples for the replicated condition, but the variability between these two and the one for the other condition would be untouched. Thus fluctuations in sampling and cooking would allow your judges to identify more than 1/3 of all cases correctly when the effects of the two storage conditions were the same. The extent of this increase would depend on the amount of variability. If you want to be sure that differences in conditions in storage are the cause of significantly more than one-third correct identification of duplicates, you must not mix the replicated peas, but instead include a sub-sample of each 300 gram sample in each set of three subsamples for judgment.

As pointed out in the answer to your second question, the amount of variability determines the false excess over one-third. If the variability were different at the two temperatures, then we would expect some effect of the sort you describe.

Even if you did not mix the replicates, differences in variability might cause differences in the apparent ability of your judges depending on which temperature was replicated. I am afraid that you should analyze the results separately for replication at the two temperatures, at least unless and until you find consistent results.

JOHN W. TUKEY

ABSTRACTS

THE BIOMETRIC SOCIETY, BRITISH REGION

Abstracts of papers for meeting on

Thursday, 8th December 1949

83 F. J. ANSCOMBE: **Theoretical distributions intended to represent population counts**

The negative binomial distribution arises as the solution to two distinct sampling problems (Feller, *Ann. Math. Stat., 14*, 1943, 389), namely (1) sampling a Poisson distribution whose mean varies randomly from observation to observation in a Type III distribution (Greenwood & Yule, *J. R. Stat. Soc., 83*, 1920, 255), (2) sampling a growing population or an aggregate of colonies, under certain conditions (Kendall, *Biometrika, 35*, 1948, 6, and Quenouille, *Biometrics, 5*, 1949, 162). If the precise assumptions in these derivations are altered, distributions other than the negative binomial result. Six other two-parameter distributions have been suggested by various authors as possibly representative of certain kinds of observations. It is found that they can be arranged in order of increasing skewness and kurtosis: (1) distribution proposed by Thomas (*Biometrika, 36*, 1949, 18), (2) Type A distribution of Neyman (*Ann. Math. Stat., 10*, 1939, 35), (3) Type B (loc. cit.), (4) Type C (loc. cit.), (5) distribution due to Aeppli and given by Polya (*Ann. Inst. Poincaré, 1*, 1930, 117), (6) the negative binomial, (7) a discrete form of the lognormal distribution (Preston, *Ecology, 29*, 1948, 254). Distributions (1) and (2) may have several modes; (5) has one or two modes; (6) and (7) have one mode. Comparisons made by Beall (*Ecology, 21*, 1940, 460) are considered. Tests are available for discriminating experimentally between the different forms of distribution.

84 J. W. BOAG: **Estimation of cure rate in cancer therapy**

Type of data available. Common measures of the success of treat-

first, but I should want to see data from many assays before expressing a final opinion.

D. J. FINNEY

76 **QUERY:** Another point which came up in connection with frozen peas (Query 74, December, 1949) bothers me. As explained there, judges were given three dishes of peas, two from one temperature and one from another. We randomized the choice of the temperature to be duplicated. On separate examination we find that the judges score more correctly during the early weeks when the lower temperature was duplicated. This has also shown up in other experiments. Can you offer any suggestion as to its cause?

We cooked three 300 gram samples of peas of each lot at each tasting period, two from one temperature and one from another. Following cooking, the two samples from one temperature were mixed before taking samples for the judges. Could this procedure have influenced results?

ANSWER: Let me take your last question first, since the answer is "yes". Suppose that the average quality and the variability from sample to sample, including the effect of cooking, were the same for peas stored under condition *A* and condition *B*. After mixing there should be almost no difference between the two samples for the replicated condition, but the variability between these two and the one for the other condition would be untouched. Thus fluctuations in sampling and cooking would allow your judges to identify more than 1/3 of all cases correctly when the effects of the two storage conditions were the same. The extent of this increase would depend on the amount of variability. If you want to be sure that differences in conditions in storage are the cause of significantly more than one-third correct identification of duplicates, you must not mix the replicated peas, but instead include a sub-sample of each 300 gram sample in each set of three subsamples for judgment.

As pointed out in the answer to your second question, the amount of variability determines the false excess over one-third. If the variability were different at the two temperatures, then we would expect some effect of the sort you describe.

Even if you did not mix the replicates, differences in variability might cause differences in the apparent ability of your judges depending on which temperature was replicated. I am afraid that you should analyze the results separately for replication at the two temperatures, at least unless and until you find consistent results.

JOHN W. TUKEY

ABSTRACTS

THE BIOMETRIC SOCIETY, BRITISH REGION

Abstracts of papers for meeting on

Thursday, 8th December 1949

83 F. J. ANSCOMBE: **Theoretical distributions intended to represent population counts**

The negative binomial distribution arises as the solution to two distinct sampling problems (Feller, *Ann. Math. Stat.*, *14*, 1943, 389), namely (1) sampling a Poisson distribution whose mean varies randomly from observation to observation in a Type III distribution (Greenwood & Yule, *J. R. Stat. Soc.*, *83*, 1920, 255), (2) sampling a growing population or an aggregate of colonies, under certain conditions (Kendall, *Biometrika*, *35*, 1948, 6, and Quenouille, *Biometrics*, *5*, 1949, 162). If the precise assumptions in these derivations are altered, distributions other than the negative binomial result. Six other two-parameter distributions have been suggested by various authors as possibly representative of certain kinds of observations. It is found that they can be arranged in order of increasing skewness and kurtosis: (1) distribution proposed by Thomas (*Biometrika*, *36*, 1949, 18), (2) Type A distribution of Neyman (*Ann. Math. Stat.*, *10*, 1939, 35), (3) Type B (loc. cit.), (4) Type C (loc. cit.), (5) distribution due to Aeppli and given by Polya (*Ann. Inst. Poincaré*, *1*, 1930, 117), (6) the negative binomial, (7) a discrete form of the lognormal distribution (Preston, *Ecology*, *29*, 1948, 254). Distributions (1) and (2) may have several modes; (5) has one or two modes; (6) and (7) have one mode. Comparisons made by Beall (*Ecology*, *21*, 1940, 460) are considered. Tests are available for discriminating experimentally between the different forms of distribution.

84 J. W. BOAG: **Estimation of cure rate in cancer therapy**

Type of data available. Common measures of the success of treat-

ment. Distribution of survival times of patients not permanently cured is approximately lognormal (parameters μ, σ). If a proportion c is cured then μ, σ and c may be estimated by maximum likelihood from the data for any sample of cases. Solution of the maximum likelihood equations by numerical iteration. σ approximately constant and can frequently be postulated in advance, thus simplifying the determination of μ and c. Examples of the use of this method. Amount of information (I_c) about c. Variation of I_c with number of cases, length of observation time and value of c. Planning a clinical experiment. Level of significance to be adopted. Statistical model of a clinical experiment.

85 A. R. G. OWEN: **Rappoport's theory of peck order.**

Classification of types of social structure as defined by peck order may be simplified by the concept of caste. The numbers of structurally distinct castes of n birds supply a basis for calculating the numbers of distinct social structures to be composed of n birds. Rappoport's classification of social structure by a partition is, however, insufficient to distinguish all isomorphic types, and examples are given of two or three types representable by the same partition.

86 E. REEVE and F. W. ROBERTSON: **An attempt to analyse the genetic variance of a strain of drosophila melanogaster, selected for long wings.**

This strain, selected continuously for 70 generations, has hardly advanced since generation 20, but shows consistently high phenotype variance for wing length and retains considerable genetic variability, since wing length and its variance decline under mass-mating.

Progeny tests were made on generations 45, 51 and 68 in order to analyse this genetic variance and to study the factors preventing further response to selection. These show that heritability of wing length is about 50% (compared with about 30% in unselected stock), and that the phenotypic variance of wing length increases with mean wing length. By taking one test to a second generation, it was possible to show that heritability of wing length increases with mean wing length of the families in which it is measured, thus suggesting that the largest parent flies are the most heterozygous for genes affecting wing length.

The main factor limiting further advance in the strain seems to be recessive lethal effects of the long-wing genes, but the existence of overdominance cannot be ruled out.